

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-P5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:27:33 ; Search time 35.0274 Seconds  
(without alignments)  
2434.679 Million cell updates/sec

Title: US-09-920-954-2

Sequence: 1 MRXKKVFLVLSNAALIST.....EVOATNPVGQFSLATVN 640

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: A\_Geneseq\_101002.\*  
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3029	97.6	640	AAV17090	Bacillus alkaline
2	3028	97.6	640	AAV17091	Bacillus alkaline
3	3020	97.4	639	AAV17089	Bacillus alkaline
4	3016	97.2	640	AAV17088	Bacillus alkaline
5	3007	96.9	639	AAV17087	An alkaline protea
6	2722.5	87.8	641	AAW89547	Bacillus JP170 pro
7	2414	77.8	636	AAW89548	Bacillus sp. alkai
8	2155	69.5	434	AAW50080	Bacillus sp. KSM-KP
9	2155	69.5	434	AAW50081	Bacillus sp. KSM-KP
10	2082	67.1	434	AAW50085	Bacillus sp alkali

11	2060.5	66.4	433	AAW50086	Bacillus sp alkali
12	1994	64.3	434	AAW50090	Bacillus sp KSM-KP
13	1952.5	62.9	433	AAW50084	Bacillus sp SD-521
14	1948.5	62.8	433	AAW50082	Bacillus sp D6-(FE
15	1941.5	62.6	433	AAW50083	Bacillus sp Y-(FER
16	1940.5	62.6	433	AAW50083	Alkali-protease ya
17	1940.5	62.6	433	AAW50083	Modified Bacillus
18	1940.5	62.6	433	AAW50083	Bacillus sp. Lion
19	1940.5	62.6	433	AAW50083	Amino acid sequenc
20	1940.5	62.6	433	AAW50083	Bacillus Lacey en
21	1644.5	53.0	345	AAW62230	Subtilase JP170. Ir
22	1644.5	53.0	345	AAW62230	Subtilase JP170. Ir
23	451.5	14.6	639	AAW24121	Thermococcus prote
24	451.5	14.6	639	AAW24121	Thermococcus prote
25	408	13.2	654	AAW24129	Pyrococcus furiosu
26	408	13.2	654	AAW24129	Hyperthermostable
27	391	12.6	659	AAW24123	Protease. Synthet
28	391	12.6	659	AAW24123	Hyperthermostable
29	391	12.6	659	AAW24123	Pyrococcus furiosu
30	391	12.6	659	AAW24123	Hyperthermostable
31	377.5	12.2	545	AAW24122	Transglutaminase r
32	358	11.5	1079	AAW24122	T. yonsei subtili
33	340	11.0	520	AAW24122	Fragment of dnpa g
34	340	11.0	520	AAW24122	Streptomyces vliid
35	340	11.0	520	AAW24122	Dnpa-mel chimeric
36	307.5	9.9	418	AAW24122	Bacillus sp. T9145
37	299.5	9.7	903	AAW24122	Hyperthermostable
38	299.5	9.7	1398	AAW24122	Protease. Pyroco
39	299.5	9.7	1398	AAW24122	Pyrococcus furiosu
40	299.5	9.7	1398	AAW24122	Pyrococcus furiosu
41	292.2	9.4	699	AAW24122	Pyrococcus furiosu
42	288.5	9.3	227	AAW24122	F. balustinus CP70
43	288.5	9.3	227	AAW24122	Hyperthermostable
44	283.5	9.2	806	AAW24122	Rp-III residual pr
45	283.5	9.1	188	AAW24122	Thermococcus prote

#### ALIGNMENTS

RESULT 1  
AAV17090 standard; Protein: 640 AA.  
ID AAV17090;  
AC AAV17090;  
XX 21-JUL-1999 (first entry)  
DT XX  
DE Bacillus alkaline protease.  
XX XX  
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;  
KW washing composition; oxidising agent.  
XX OS Bacillus sp.  
XX FN W0918218-A1.  
XX PD 15-APR-1999.  
XX PF 07-OCT-1998; 98WO-JP04528.  
XX PR 07-OCT-1997; 97JP-0274570.  
XX (KAOS) KAO CORP.  
XX Hitomi J., Kageyama Y., Kubota H., Nomura M.,  
XX Sasaki K., Shikata S., Takaiwa M;  
XX WPI: 1999-287736/24.  
XX N-PSDB; AA37278.  
XX Alkali protease from Bacillus used in washing powders  
XX



QY 361 TILFAENEXPNGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420  
 DB 361 TILFAENEXPNGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420  
 QY 421 RIKPDVAPCTXILISARSSILAPDSSFFMANHDSKYAVNGCTSMATPIVAGNVAQLREHFVK 480  
 DB 421 RIKPDVAPCTXILISARSSILAPDSSFFMANHDSKYAVNGCTSMATPIVAGNVAQLREHFVK 480  
 QY 481 NRGITPPKSLILKAALILAGAADIGLGYPNGNGMGRTLDKSLNVAAYVNESSLSLSOKAT 540  
 DB 481 NRGITPPKSLILKAALILAGAADIGLGYPNGNGMGRTLDKSLNVAAYVNESSLSLSOKAT 540  
 QY 541 YXFATAGKFLKISLVMSDAPASTASVTLVNDLVLITAPNCTXYVGNDEFPXKXMMMD 600  
 DB 541 YXFATAGKFLKISLVMSDAPASTASVTLVNDLVLITAPNCTXYVGNDEFPXKXMMMD 600  
 QY 601 GRNNVEVFNNXQSGTYTIEVOAYNVPGQFSLAIYN 640  
 DB 601 GRNNVEVFNNXQSGTYTIEVOAYNVPGQFSLAIYN 640

RESULT 3  
 AAY17089 standard; protein: 639 AA.  
 ID AAY17089;  
 AC AAY17089;  
 XX 21-JUL-1999 (first entry)  
 DT 21-JUL-1999 (first entry)  
 DE Bacillus alkaline protease.  
 KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;  
 KM washing composition; oxidizing agent.  
 OS Bacillus sp.  
 PN WO9918218-A1.  
 PD 15-APR-1999.  
 PF 07-OCT-1998; 98WO-JP04528.  
 PR 07-OCT-1997; 97JP-0274570.  
 PA (KAO) KAO CORP.  
 PI Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;  
 PI Saeki K, Shikata S, Takaiwa M;  
 DR WPI: 1999-287736/24.  
 DR N-PSDB: AAX37277.  
 PT Alkali protease from Bacillus used in washing powders  
 PS Disclosure: Page 53-58; 71pp; Japanese.  
 CC The invention relates to alkaline proteases produced by strains of  
 CC Bacillus. The proteases ability to digest casein is not inhibited by  
 CC oleic acid and they have a high stability to oxidizing agents. The  
 CC alkaline protease of the invention has the following properties: (a) it  
 CC is active over the pH range 4-13 and has at least 80% of its optimum  
 CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is  
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)  
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has  
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be  
 CC used as enzymes in washing compositions for use in automatic dishwashers  
 CC and for washing clothes. The stability to oxidizing agents allows the  
 CC enzyme to be an effective component of washing compositions including  
 CC bleaches. The present sequence represents an alkaline protease.  
 SO Sequence 639 AA;  
 Query Match 97.4%; Score 3020; DB 20; Length 639.

Best Local Similarity 93.4%; Pred. No. 2,8e-233;  
 Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;  
 QY 4 KKKVFLSVLSAAAILSTVALKNPSAGAKRFFDLDFKGIOTTTDXGFSKOKOTGAALFLV 63  
 DB 3 KKKVFLSVLSAAAILSTVALKNPSAGAKRFFDLDFKGIOTTTDXGFSKOKOTGAALFLV 62  
 QY 64 ESENVKLLXGKIKLETPVANNKLIHQFNGPILTEFKOXLEXTGAKTILDIPTPYATIVE 123  
 DB 64 ESENVKLLXGKIKLETPVANNKLIHQFNGPILTEFKOXLEXTGAKTILDIPTPYATIVE 122  
 QY 124 YEGDVYXXXHXIEHVESVPEPLPYXIIDPOLFTKGSXYKAXALDROKXEVOLGIE 163  
 DB 124 YEGDVYXXXHXIEHVESVPEPLPYXIIDPOLFTKGSXYKAXALDROKXEVOLGIE 162  
 QY 184 XIAQXXXNDVXYITAKPEYKVMNDVARGYKADVAOSSTGLGGOQIYAAVADGLDTR 243  
 DB 184 XIAQXXXNDVXYITAKPEYKVMNDVARGYKADVAOSSTGLGGOQIYAAVADGLDTR 242  
 QY 244 NDSMHEAFRGKITLALYALGRTNANDTNGHGHVAGSVLGNGXTNKGMAPQANLVFQSI 303  
 DB 244 NDSMHEAFRGKITLALYALGRTNANDTNGHGHVAGSVLGNGXTNKGMAPQANLVFQSI 302  
 QY 304 MDSXGIGLGLPSNLQTLFSSQASAGARLHTNSGAAVNGAYTTDSRNVDYVRKNDMTIL 363  
 DB 304 MDSXGIGLGLPSNLQTLFSSQASAGARLHTNSGAAVNGAYTTDSRNVDYVRKNDMTIL 362  
 QY 364 FAAGNEXPNGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIK 423  
 DB 364 FAAGNEXPNGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIK 422  
 QY 424 PDVWAGTXYLSNRSSILAPDSSFFMANHDSKYAVNGCTSMATPIVAGNVAQLREHFVNRG 483  
 DB 424 PDVWAGTXYLSNRSSILAPDSSFFMANHDSKYAVNGCTSMATPIVAGNVAQLREHFVNRG 482  
 QY 484 ITRKPSILKAALILAGAADIGLGYPNGNGMGRTLDKSLNVAAYVNESSLSLSOKATYXP 543  
 DB 484 ITRKPSILKAALILAGAADIGLGYPNGNGMGRTLDKSLNVAAYVNESSLSLSOKATYXP 542  
 QY 544 TATAGKFLKISLVMSDAPASTASVTLVNDLVLITAPNCTXYVGNDEFPXKXMMMDGRN 603  
 DB 544 TATAGKFLKISLVMSDAPASTASVTLVNDLVLITAPNCTXYVGNDEFPXKXMMMDGRN 602  
 QY 604 NVERVFNNXQSGTYTIEVOAYNVPGQFSLAIYN 640  
 DB 604 NVERVFNNXQSGTYTIEVOAYNVPGQFSLAIYN 639

RESULT 4  
 AAY17088 standard; protein: 640 AA.  
 ID AAY17088;  
 AC AAY17088;  
 XX 21-JUL-1999 (first entry)  
 DT 21-JUL-1999 (first entry)  
 DE An alkaline protease sequence from Bacillus species.  
 KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;  
 KM washing composition; oxidizing agent.  
 OS Bacillus sp.  
 PN WO9918218-A1.  
 PD 15-APR-1999.  
 PF 07-OCT-1998; 98WO-JP04528.  
 FT Key location/Qualifiers  
 FT Misc-difference 1..640  
 FT "all residues indicated as Xaa are arbitrary  
 FT amino acids"

```

xx 07-OCT-1997; 97JP-0274570.
PR (KAOS ) KAO CORP.
PA
xx Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;
PI Saeki K, Shikata S, Takaiwa M;
xx
DR WPI: 1999-287736/24.
DR N-PSDB: AAX37278.
xx
xx Alkali protease from Bacillus used in washing powders
xx
xx Claim 3; Page 50-53; 71pp; Japanese.
xx
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease of the
CC invention.
xx
SQ Sequence 640 AA;
xx
Query Match 97.2%; Score 3016; DB 20; Length 640;
Best Local Similarity 100.0%; Pred. No. 5.9e-233;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
xx
QY 1 MEXKKRVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIOTTTDXXGFSKQXGTGA 60
DB 1 MEXKKRVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIOTTTDXXGFSKQXGTGA 60
QY 61 FLVESNNVNLKGLKXKLETFVPANNKLTIXQFNPILEETKQXLETKAKIIDYIPDAY 120
DB 61 FLVESNNVNLKGLKXKLETFVPANNKLTIXQFNPILEETKQXLETKAKIIDYIPDAY 120
QY 121 IYEYEDVAXXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVKAKALDTKQXNKREJOR 180
DB 121 IYEYEDVAXXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVKAKALDTKQXNKREJOR 180
QY 181 GIEIXIAQXXXSNDVXYITAKPEKRYVNDVARGIVKADVAQSSYGLYGQGIYAVADTGLD 240
DB 181 GIEIXIAQXXXSNDVXYITAKPEKRYVNDVARGIVKADVAQSSYGLYGQGIYAVADTGLD 240
QY 241 TERNDSMEHARFGKLTALYALGRTNANDTNGHGHVAGSVLGNKXTKKGAPOANLVF 300
DB 241 TERNDSMEHARFGKLTALYALGRTNANDTNGHGHVAGSVLGNKXTKKGAPOANLVF 300
QY 301 OSIMDSXGGLGSLPNTQLFESQASAGARHTNSGAAVNGAYTTDSRNVDYVRKNDM 360
DB 301 OSIMDSXGGLGSLPNTQLFESQASAGARHTNSGAAVNGAYTTDSRNVDYVRKNDM 360
QY 361 TILFPAAGNEXPNMGITISAGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
DB 361 TILFPAAGNEXPNMGITISAGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
QY 421 RIKPVMAGTYILSARSSLAPDSSFWANHDSKYATMGTSNATPIVAQNVQLRHFVK 480
DB 421 RIKPVMAGTYILSARSSLAPDSSFWANHDSKYATMGTSNATPIVAQNVQLRHFVK 480
QY 481 NRGITPKEPLLKAAALAGAADXGLGYPNGOGMGRVTLTKSLNVAAVNSSXLSSTQKAT 540
DB 481 NRGITPKEPLLKAAALAGAADXGLGYPNGOGMGRVTLTKSLNVAAVNSSXLSSTQKAT 540
QY 541 YXFTATAGKPLKISLWMSDAPASTASTAVTLVNDLVLITAPNGTXVGNDFXXPXXXND 600
DB 541 YXFTATAGKPLKISLWMSDAPASTASTAVTLVNDLVLITAPNGTXVGNDFXXPXXXND 600

```

```

DB 541 YXFTATAGKPLKISLWMSDAPASTASTAVTLVNDLVLITAPNGTXVGNDFXXPXXXND 600
QY 601 GRNNVENFINKPQSGTYTIEVOAYNVPGPOXFESLAIYN 640
DB 601 GRNNVENFINKPQSGTYTIEVOAYNVPGPOXFESLAIYN 640
xx
xx RESULT 5
xx AAY17087
xx ID AAY17087 standard; protein; 639 AA.
xx
xx AC AAY17087;
xx
xx DT 21-JUL-1999 (first entry)
xx
xx DE An alkaline protease sequence from Bacillus species.
xx
xx KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
xx washing composition; oxidising agent.
xx
xx OS Bacillus sp.
xx
xx FH Key Location/Qualifiers
xx FT Misc-difference 1..639
xx FT "all residues indicated as Xaa are arbitrary
xx FT amino acids"
xx
xx PN W09918218-1.
xx
xx PD 15-APR-1999.
xx
xx PF 07-OCT-1998; 98WO-JP04528.
xx
xx PR 07-OCT-1997; 97JP-0274570.
xx
xx PA (KAOS ) KAO CORP.
xx
xx PI Hitomi J, Kageyama Y, Kubota H, Nomura M, Okuda M;
xx PI Saeki K, Shikata S, Takaiwa M;
xx
xx DR WPI: 1999-287736/24.
xx DR N-PSDB: AAX37277.
xx
xx PT Alkali protease from Bacillus used in washing powders
xx
xx PS Claim 3; Page 47-50; 71pp; Japanese.
xx
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease of the
CC invention.
xx
SQ Sequence 639 AA;
xx
Query Match 96.9%; Score 3007; DB 20; Length 639;
Best Local Similarity 100.0%; Pred. No. 3.1e-232;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
xx
QY 4 KKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIOTTTDXXGFSKQXGTGA 63
DB 3 KKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIOTTTDXXGFSKQXGTGA 62

```

```

QY 64 ESENVKLGKXKKLETPANNKLIHQFNGPILEETKQXLEXTGAKILIDYIPDYATVE 123
DB 63 ESENVKLGKXKKLETPANNKLIHQFNGPILEETKQXLEXTGAKILIDYIPDYATVE 122
QY 124 YEGDVXSKXXIIEHVESVEPYLPYXIDPOLFTKGASXLVKAXALDTRQXKKEVOLRGIE 183
DB 123 YEGDVXSKXXIIEHVESVEPYLPYXIDPOLFTKGASXLVKAXALDTRQXKKEVOLRGIE 182
QY 184 XIAQXXSNDVYXITAKPEYKVMNDVARGIVKADVQSSGYGOGQIVAAVADTGLDTRG 243
DB 183 XIAQXXSNDVYXITAKPEYKVMNDVARGIVKADVQSSGYGOGQIVAAVADTGLDTRG 242
QY 244 NDSMHEAFRGKITLALYALGRTNANDTNGHGVHAGSVLGGXNKGMAPQANLVEQST 303
DB 243 NDSMHEAFRGKITLALYALGRTNANDTNGHGVHAGSVLGGXNKGMAPQANLVEQST 302
QY 304 MDXSGGLGGLPSNLQTLFSGQXSGARLHTNSGCAVNGAYTTDSRNVDDYVRKNDMTIL 363
DB 303 MDXSGGLGGLPSNLQTLFSGQXSGARLHTNSGCAVNGAYTTDSRNVDDYVRKNDMTIL 362
QY 364 PAAGNEXPGGTTISAPGTAKNAITVGATEMLRPSFGSYADNINHVAQFSSRGPTKGRIRK 423
DB 363 PAAGNEXPGGTTISAPGTAKNAITVGATEMLRPSFGSYADNINHVAQFSSRGPTKGRIRK 422
QY 424 PDVWABGTIILSARSSSLAPDSSFWANHDSKYAYMGTSNATPIVAGNVAOLREHFVNRG 483
DB 423 PDVWABGTIILSARSSSLAPDSSFWANHDSKYAYMGTSNATPIVAGNVAOLREHFVNRG 482
QY 484 ITPKPSILKAALITAGAAIXGLGYFNGNGRGRTLDKSLNVAVYVNESSXLSSTQKATYXF 543
DB 483 ITPKPSILKAALITAGAAIXGLGYFNGNGRGRTLDKSLNVAVYVNESSXLSSTQKATYXF 542
QY 544 TATAGKPLKISLWSDAPASTTASVTLVNDLDVITAPNGTYXVGNDRXXKXNMNGRN 603
DB 543 TATAGKPLKISLWSDAPASTTASVTLVNDLDVITAPNGTYXVGNDRXXKXNMNGRN 602
QY 604 NVENFINKPGSGTYTIEVQAINVPGPQXFSIAIVN 640
DB 603 NVENFINKPGSGTYTIEVQAINVPGPQXFSIAIVN 639

```

RESULT 6  
AAM89547  
ID AAM89547 standard; Protein: 641 AA.  
AC AAM89547;  
XX  
DT 12-APR-1999 (first entry)  
DE Bacillus jpl70 protease.  
XX  
KW Protease; detergent; surfactant; leather processing; debittering;  
KW flavour.  
XX  
OS Bacillus sp.  
XX  
FH Key  
FT Peptide  
FT Region  
FT Protein  
FT  
FT  
XX MO9856927-A2.  
XX 17-DEC-1998.  
XX  
XX 09-JUN-1998: 98WO-US12005.  
XX  
XX 12-JUN-1997: 97US-0873479.  
XX  
XX (NOVO ) NOVO NORDISK BIOTECH INC.

```

XX XX Christlanson L, Sloma A:
PI PT
XX DR MPI: 1999-080908/07.
XX DR N-PSDB: AAV82382.
PT PT Novel protease from Bacillus subtilis LC20 - useful in laundry and
XX PT dishwashing detergents and for leather processing
XX PS Claim 7: Page 53-54; 77pp: English.
XX
CC This is the amino acid sequence of a novel protease of Bacillus sp.
CC jpl70 (NCBI 12513), as deduced from the nucleotide sequence of an
CC isolated gene (see AAV82382). The entire protein, including the
CC signal peptide and prepro region, has 778 identity to alkaline
CC protease I (see AAM89548) from Bacillus. The invention provides
CC vectors, recombinant host cells and methods for the recombinant
CC production of the protease. The protease is used in laundry and
CC dishwashing detergents, for institutional and industrial cleaning,
CC and for leather processing, as well as for debittering and
CC enhancing the degree of hydrolysis of protein hydrolysates, for
CC flavour development through hydrolysis of proteins, degradation of
CC undigested peptides and in enzymatic synthesis of peptides. It has
CC enhanced stability towards oxidation under alkaline conditions,
CC e.g. towards bleaching agents of the peroxy type. The invention
CC also provides mutant cells in which the protease activity is
CC diminished. Such cells can be used for the production of
CC heterologous recombinant proteins.
XX
SO Sequence 641 AA:

```

Query Match 87.8%; Score 2722.5; DB 20: Length 641;  
Best Local Similarity 82.9%; Pred. No. 1.9e-209;  
Matches 532; Conservative 40; Mismatches 67; Indels 3; Gaps 2;

```

QY 1 MRKK--KRVLSLVSAAILSTVALNPSAGARXFDLDFKGIOTTTDXGFSKXQGTGA 58
DB 1 MRKKSQRVFLSVLSVALLSVALSPSTIGANNFELDFGIEFLTLERAAKRGKGTGK 60
QY 59 AAFVSEENVKLGKXKKLETPANNKLIHQFNGPILEETKQXLEXTGAKILIDYIPDY 118
DB 61 ASPLVSENVKIPKSIQKLEVPADNKLIVQFDGPILEETQQLQLEKGTAKILIDYIPDY 120
QY 119 AYIVEEGDVXSKXXIIEHVESVEPYLPYXIDPOLFTKGASXLVKAXALDTRQXKEVO 178
DB 121 AYIVEEDGVKAYVNAIHALESVEPYLPYXIDPOLFTKGASXLVKAXALDTRQXKEVR 180
QY 179 LAGIEIXAQXXSNDVYXITAKPEYKVMNDVARGIVKADVQSSGYGOGQIVAAVADTG 238
DB 181 LKLEQIAQYATNDVLYVTPKPEYELVNDVARGIVKADVQNNFGLGGQIVAAVADTG 240
QY 239 LDTGRNDSMHEAFRGKITLALYALGRTNANDTNGHGVHAGSVLGGXNKGMAPQANL 298
DB 241 LDTGRNDSMHEAFRGKITLALYALGRTNANDTNGHGVHAGSVLGN-ATNKGMAPQANL 299
QY 299 VFQSIMDSXGGLGGLPSNLQTLFSGQXSGARLHTNSGCAVNGAYTTDSRNVDDYVRKN 358
DB 300 VFQSIMDSXGGLGGLPANLQTLFSGQYVSAGARLHTNSGCAVNGAYTTDSRNVDDYVRKN 359
QY 359 DMTILFAAGNEXPGGTTISAPGTAKNAITVGATEMLRPSFGSYADNINHVAQFSSGPTK 418
DB 360 DMTILFAAGNEXPGGTTISAPGTAKNAITVGATEMLRPSFGSYADNINHVAQFSSGPTK 419
QY 419 DGRIRPDVWABGTIILSARSSSLAPDSSFWANHDSKYAYMGTSNATPIVAGNVAOLREHF 478
DB 420 DGRIRPDVWABGTIILSARSSSLAPDSSFWANHDSKYAYMGTSNATPIVAGNVAOLREHF 479
QY 479 VNRKGTTRKPSILKAALITAGAAIXGLGYFNGNGRGRTLDKSLNVAVYVNESSXLSSTQK 538
DB 480 VNRKGTTRKPSILKAALITAGAAIXGLGYFNGNGRGRTLDKSLNVAVYVNESSXLSSTQK 539
QY 539 ATYXFTATAGKPLKISLWSDAPASTTASVTLVNDLDVITAPNGTYXVGNDRXXKXNMNGRN 598

```





CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This  
 CC sequence represents a fragment of the alkaline protease Kp43 from  
 CC *Bacillus* sp strain KSM-KP43 which is used to create the modified protease  
 CC represented in AAM50090.

XX Sequence 434 AA:

Query Match 69.5%: Score 2155; DB 23: Length 434;  
 Best Local Similarity 96.3%: Pred. No. 3,1e-164;  
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

QY 207 NDVARGIVADVAQSSYGLYGQGIYVAVDTGLDTGRNDSMHEAFRGKITALYALGRTN 266
    1 NDVARGIVADVAQSSYGLYGQGIYVAVDTGLDTGRNDSMHEAFRGKITALYALGRTN 60
DB 267 NANDTNGHGTIVAGSVLGNKXNKGMAPQANLVFOSIMDSXGGLGSPNLQTLFSGQAXS 326
    61 NANDTNGHGTIVAGSVLGNKXNKGMAPQANLVFOSIMDSXGGLGSPNLQTLFSGQAXS 120
QY 327 AGARITNSMGAAYNGAYTTDSRVNDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 386
    121 AGARITNSMGAAYNGAYTTDSRVNDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
DB 387 TVGATENLRPSFGSYADNININVAQFSSRGPTKDGRIKPDVMAPGTYIISARSLAPDSSF 446
    181 TVGATENLRPSFGSYADNININVAQFSSRGPTKDGRIKPDVMAPGTYIISARSLAPDSSF 240
QY 447 WANHDSKYAYMGSTMATPPIVAGNVAQLREHFVKNKGITPKPSLLKALLIAGADVGLGY 506
    241 WANHDSKYAYMGSTMATPPIVAGNVAQLREHFVKNKGITPKPSLLKALLIAGADVGLGY 300
DB 507 PNGNQGMRVTLDDKSLNVAAYNESSXLSSTOKATYFTATAGKPLKISLWSDAPASTTA 566
    301 PNGNQGMRVTLDDKSLNVAAYNESSXLSSTOKATYFTATAGKPLKISLWSDAPASTTA 360
QY 567 SVTLVNDLDELVTAPNGTXYGNDFFXPPXXXNMDGRNNEVFINKPQSGTYTTEVOAYN 626
    361 SVTLVNDLDELVTAPNGTXYGNDFFXPPXXXNMDGRNNEVFINKPQSGTYTTEVOAYN 420
DB 627 VPGPQXPSLAIYN 640
    421 VPGPQXPSLAIYN 434
  
```

RESULT 9  
 AAM50081  
 ID AAM50081 standard; protein: 434 AA.

XX AAM50081;  
 AC 12-AUG-2002 (first entry)  
 DT 12-AUG-2002 (first entry)  
 XX *Bacillus* sp KSM-KP9860 alkaline protease protein fragment.  
 DE Alkaline protease; detergent; laundry; bleaching; dishwasher.  
 KW *Bacillus* sp.  
 OS EPI209233-A2.  
 XX EPI209233-A2.  
 PD 29-MAY-2002.  
 XX 22-NOV-2001; 2001EP-0127851.  
 PF 22-NOV-2000; 2000JP-0355166.  
 PR 12-APR-2001; 2001JP-0114048.  
 XX (KAOS ) KAO CORP.  
 PA Hatada Y., Ogawa A., Kageyama Y., Sato T., Araki H., Sumitomo N.,  
 PI Okuda M., Saeki K.,  
 XX WPI, 2002-437518/47.

XX New modified alkaline proteases useful in detergent compositions -  
 PT Claim 5; Page 12-13; 25pp; English.

CC This invention describes novel *Bacillus* sp. alkaline proteases useful in  
 CC detergent compositions, especially in laundry, bleaching or automatic  
 CC dishwasher detergents. The novel proteases have an increased detergency %  
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This  
 CC sequence represents a fragment of the alkaline protease Kp9860 from  
 CC *Bacillus* sp strain KSM-KP9860 described in the method of the invention.

XX Sequence 434 AA:

Query Match 69.5%: Score 2155; DB 23: Length 434;  
 Best Local Similarity 96.3%: Pred. No. 3,1e-164;  
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

QY 207 NDVARGIVADVAQSSYGLYGQGIYVAVDTGLDTGRNDSMHEAFRGKITALYALGRTN 266
    1 NDVARGIVADVAQSSYGLYGQGIYVAVDTGLDTGRNDSMHEAFRGKITALYALGRTN 60
DB 267 NANDTNGHGTIVAGSVLGNKXNKGMAPQANLVFOSIMDSXGGLGSPNLQTLFSGQAXS 326
    61 NANDTNGHGTIVAGSVLGNKXNKGMAPQANLVFOSIMDSXGGLGSPNLQTLFSGQAXS 120
QY 327 AGARITNSMGAAYNGAYTTDSRVNDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 386
    121 AGARITNSMGAAYNGAYTTDSRVNDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
DB 387 TVGATENLRPSFGSYADNININVAQFSSRGPTKDGRIKPDVMAPGTYIISARSLAPDSSF 446
    181 TVGATENLRPSFGSYADNININVAQFSSRGPTKDGRIKPDVMAPGTYIISARSLAPDSSF 240
QY 447 WANHDSKYAYMGSTMATPPIVAGNVAQLREHFVKNKGITPKPSLLKALLIAGADVGLGY 506
    241 WANHDSKYAYMGSTMATPPIVAGNVAQLREHFVKNKGITPKPSLLKALLIAGADVGLGY 300
DB 507 PNGNQGMRVTLDDKSLNVAAYNESSXLSSTOKATYFTATAGKPLKISLWSDAPASTTA 566
    301 PNGNQGMRVTLDDKSLNVAAYNESSXLSSTOKATYFTATAGKPLKISLWSDAPASTTA 360
QY 567 SVTLVNDLDELVTAPNGTXYGNDFFXPPXXXNMDGRNNEVFINKPQSGTYTTEVOAYN 626
    361 SVTLVNDLDELVTAPNGTXYGNDFFXPPXXXNMDGRNNEVFINKPQSGTYTTEVOAYN 420
DB 627 VPGPQXPSLAIYN 640
    421 VPGPQXPSLAIYN 434
  
```

RESULT 10  
 AAM50085  
 ID AAM50085 standard; protein: 434 AA.

XX AAM50085;  
 AC 12-AUG-2002 (first entry)  
 DT 12-AUG-2002 (first entry)  
 XX *Bacillus* sp alkaline protease protein A-1 fragment.  
 DE Alkaline protease; detergent; laundry; bleaching; dishwasher.  
 KW *Bacillus* sp.  
 OS EPI209233-A2.  
 XX EPI209233-A2.  
 PD 29-MAY-2002.  
 XX 22-NOV-2001; 2001EP-0127851.  
 PF 22-NOV-2000; 2000JP-0355166.  
 PR 12-APR-2001; 2001JP-0114048.

XX (KAOS ) KAO CORP.  
PA Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;  
PI Okuda M, Saeki K;  
XX WPI: 2002-437518/47.  
XX  
XX New modified alkaline proteases useful in detergent compositions -  
XX  
XX Claim 5; Page 18-19; 25pp; English.  
XX  
XX This invention describes novel Bacillus sp. alkaline proteases useful in  
CC detergent compositions, especially in laundry, bleaching or automatic  
CC dishwasher detergents. The novel proteases have an increased detergency %  
CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This  
CC sequence represents a fragment of the alkaline protease A-1 from  
CC Bacillus sp NCIB12289 described in the method of the invention.  
XX  
XX Sequence 434 AA:  
SQ  
Query Match 67.1%; Score 2082; DB 23; Length 434;  
Best Local Similarity 91.5%; Pred. No. 2.1e-158;  
Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;  
QY 207 NDVARGIVKADYAGSSYGLYGQGIIVAAVDGLDTRGRDSSMHEAFRGKITALVALGRTN 266  
DB 1 NDVARGIVKADYAGSSYGLYGQGIIVAAVDGLDTRGRDSSMHEAFRGKITALVALGRTN 60  
QY 267 NANDTNGHGTTHVAGSVLNGXTNKGMAPOANLVFOSIMDSXGGLGCLPSNLOTLPFSQAXS 326  
DB 61 NANDTNGHGTTHVAGSVLNGXTNKGMAPOANLVFOSIMDSXGGLGCLPSNLOTLPFSQAXS 120  
QY 327 AGARHTNSMGAIVNAGVATYDTSRNVDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 386  
DB 121 AGARHTNSMGAIVNAGVATYDTSRNVDDYVRKNDMAVLPAAGNEGPGGTISAPGTAKNAI 180  
QY 387 TVGATENLRPSFGSYADNINNHVAFSSRGPTKDGRIKPDVMAPGTYIISARSSLAPDSSF 446  
DB 181 TVGATENLRPSFGSYADNINNHVAFSSRGPTKDGRIKPDVMAPGTYIISARSSLAPDSSF 240  
QY 447 WANHDSKYAYMGGTSMATPIVAGNVAOLREHFVKRNGITPKPSLKAALIAAGADYGLGY 506  
DB 241 WANHDSKYAYMGGTSMATPIVAGNVAOLREHFVKRNGITPKPSLKAALIAAGADYGLGY 300  
QY 507 PNGNOGMRVTLTKSLNVAAYVNESSXLSSTOKATYXFTATACKPLKISLWSDAPASTTA 566  
DB 301 PNGNOGMRVTLTKSLNVAAYVNESSXLSSTOKATYXFTATACKPLKISLWSDAPASTTA 360  
QY 567 SVTLVNDLDLVITAPNGTXYYGNDFXPPXXNMDCGRNNVENFINKXPOSGTYTIEVOAYN 626  
DB 361 SVTLVNDLDLVITAPNGTXYYGNDFXPPXXNMDCGRNNVENFINKXPOSGTYTIEVOAYN 420  
QY 627 VPVGPOXFSLAIYN 640  
DB 421 VPVGPOXFSLAIYN 434

RESULT 11  
AAM50086  
ID AAM50086 standard; protein: 433 AA.  
XX  
XX AAM50086;  
XX  
XX 12-AUG-2002 (first entry)  
XX  
XX Bacillus sp alkaline protease protein A-2 fragment.  
XX  
XX Alkaline protease; detergent; laundry; bleaching; dishwasher.  
XX  
XX Bacillus sp.  
XX  
XX EPI209233-A2.  
FN

XX  
PD 29-MAY-2002.  
XX  
XX 22-NOV-2001; 2001EP-0127851.  
XX  
XX PF  
XX 22-NOV-2000; 2000JP-0355166.  
XX  
XX PR 12-APR-2001; 2001JP-0114048.  
XX  
XX  
XX (KAOS ) KAO CORP.  
XX  
XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;  
PI Okuda M, Saeki K;  
XX  
XX WPI: 2002-437518/47.  
XX  
XX New modified alkaline proteases useful in detergent compositions -  
XX  
XX Claim 5; Page 20-21; 25pp; English.  
XX  
XX This invention describes novel Bacillus sp. alkaline proteases useful in  
CC detergent compositions, especially in laundry, bleaching or automatic  
CC dishwasher detergents. The novel proteases have an increased detergency %  
CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This  
CC sequence represents a fragment of the alkaline protease A-2 from  
CC Bacillus sp NCIB12513 described in the method of the invention.  
XX  
XX Sequence 433 AA:  
SQ  
Query Match 66.4%; Score 2060.5; DB 23; Length 433;  
Best Local Similarity 91.7%; Pred. No. 1.1e-156;  
Matches 398; Conservative 12; Mismatches 23; Indels 1; Gaps 1;  
QY 207 NDVARGIVKADYAGSSYGLYGQGIIVAAVDGLDTRGRDSSMHEAFRGKITALVALGRTN 266  
DB 1 NDVARGIVKADYAGSSYGLYGQGIIVAAVDGLDTRGRDSSMHEAFRGKITALVALGRTN 60  
QY 267 NANDTNGHGTTHVAGSVLNGXTNKGMAPOANLVFOSIMDSXGGLGCLPSNLOTLPFSQAXS 326  
DB 61 NANDTNGHGTTHVAGSVLNGXTNKGMAPOANLVFOSIMDSXGGLGCLPSNLOTLPFSQAXS 119  
QY 327 AGARHTNSMGAIVNAGVATYDTSRNVDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 386  
DB 120 AGARHTNSMGAIVNAGVATYDTSRNVDDYVRKNDMTILFAAGNEGPGGTISAPGTAKNAI 179  
QY 387 TVGATENLRPSFGSYADNINNHVAFSSRGPTKDGRIKPDVMAPGTYIISARSSLAPDSSF 446  
DB 180 TVGATENLRPSFGSYADNINNHVAFSSRGPTKDGRIKPDVMAPGTYIISARSSLAPDSSF 239  
QY 447 WANHDSKYAYMGGTSMATPIVAGNVAOLREHFVKRNGITPKPSLKAALIAAGADYGLGY 506  
DB 240 WANHDSKYAYMGGTSMATPIVAGNVAOLREHFVKRNGITPKPSLKAALIAAGADYGLGY 299  
QY 507 PNGNOGMRVTLTKSLNVAAYVNESSXLSSTOKATYXFTATACKPLKISLWSDAPASTTA 566  
DB 300 PNGNOGMRVTLTKSLNVAAYVNESSXLSSTOKATYXFTATACKPLKISLWSDAPASTTA 359  
QY 567 SVTLVNDLDLVITAPNGTXYYGNDFXPPXXNMDCGRNNVENFINKXPOSGTYTIEVOAYN 626  
DB 360 SVTLVNDLDLVITAPNGTXYYGNDFXPPXXNMDCGRNNVENFINKXPOSGTYTIEVOAYN 419  
QY 627 VPVGPOXFSLAIYN 640  
DB 420 VPVGPOXFSLAIYN 433

RESULT 12  
AAM50090  
ID AAM50090 standard; protein: 434 AA.  
XX  
XX AAM50090;  
XX  
XX 12-AUG-2002 (first entry)  
XX  
XX



DB	421	VPGQPTFSLAIVN	434
<hr/>			
RESULT 13			
ID	AAM50084	standard; protein; 433 AA.	
XX	AAM50084;		
XX			
DT	12-AUG-2002	(first entry)	
XX			
XX	Bacillus sp SD-521 (FERM BP-11162) alkaline protease protein fragment.		
XX			
KW	Alkaline protease; detergent; laundry; bleaching; dishwasher.		
XX			
OS	Bacillus sp.		
XX	EPI209233-A2.		
FN			
PD	29-MAY-2002.		
XX			
PE	22-NOV-2001; 2001EP-0127851.		
XX			
PR	22-NOV-2000; 2000JP-0355166.		
PR	12-APR-2001; 2001JP-0114048.		
XX			
PA	(KAOS ) KAO CORP.		
XX			
PL	Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;		
PI	Okuda M, Saeki K;		
XX			
DR	WPI; 2002-437518/47.		
PT	New modified alkaline proteases useful in detergent compositions -		
XX			
PS	Claim 5; Page 16-18; 25pp; English.		
XX			
CC	This invention describes novel Bacillus sp. alkaline proteases useful in		
CC	detergent compositions, especially in laundry, bleaching or automatic		
CC	dishwasher detergents. The novel proteases have an increased detergency &		
CC	(34-38%) compared to prior art alkaline proteases (31 and 23%). This		
CC	sequence represents a fragment of the alkaline protease SD-521 from		
CC	Bacillus sp strain SD-521 (FERM BP-11162) described in the method of the		
CC	invention.		
XX			
SQ	Sequence 433 AA:		
<hr/>			
Query Match	62.9%; Score 1952.5; DB 23; Length 433;		
Best Local Similarity	86.6%; Pred. No. 5e-148;		
Matches 376; Conservative	20; Mismatches 37; Indels 1; Gaps 1,		
OY	207 NDVAGIYKADVAOSSGILGYGGOGCIYAAVDRLDGRNDSSMHEAFRGKITLALYALGRVN	266	
DB	1 NDVAGIYKADVAQNNGNYGLGYGGQYVAADVLTDLTGRNDSSMHEAFRGKITLALYALGRVN	60	
OY	267 NANDNTNGHSTHVAGSVLYLGXGTNKGMAPQANLVFPOSIMDSXGLGLPENTLQLFSOAXS	326	
DB	61 NANDNTNGHSTHVAGSVLYGN -ALNKGMAPQANLVFPOSIMDSXGLGLPENTLQLFSQAWN	119	
OY	327 AGARIHTNSWGAAVNGAYTTDSRNVDVIYRKNDMTLLFAAGNEXPNGGTISAPGTAKNAI	386	
DB	120 AGARIHTNSWGAPVNGAYTAATANSROVEYVRNDDMTVLFAAGNEGPNSGTISAPGTAKNAI	179	
OY	387 TVGATEINRPSPGSYADNRNNHVAFOSSRCPTDGRDKPKPYMAPCXTIILARSLSLPDSSF	446	
DB	180 TVGATEINRPSPGSYLADNNHNIAQFSSSRATRDGRKKPVTVAPGTITILARSLSLPDSSF	239	
OY	447 WANHSKYAVMGCTSMATPIVAGANVAQLLEHHVKNKGIPKPSLKLAALITAGAADGLGX	506	
DB	240 WANYSKYAVMGCTSMATPIVAGANVAQLREHRIKRRGITPKRSLIKLAALITAGATDVGLCY	299	
OY	PNGNGGWGVTLTDKSLNVAIVNESSXLSTSQKATYYFTATAACKPLKISLWSDAPASTTA	566	

[illegible]

[illegible]

RESULT 15  
AAM50083  
ID AAM50083 standard; protein; 433 AA  
XX

AC AAM50083;  
yy

DT	12-AUG-2002	{first entry}
XY		

DE *Bacillus* sp Y-(FERM BP-1029) alkaline protease protein fragment  
XX

Alkaline protease; detergent; laundry; bleaching; dishwasher

03 BACILLUS sp.  
XX

XX  
XX

[illegible]

XX  
XX  
00-NOV-2000 0000 TR 00000000

PR 12-APR-2001; 2001JP-0114048.  
XX

PA (KAOS ) KAO CORP.  
XX

PI Hatada Y, Ugawa A, Kageyama Y, Sato T, Araki H, Sumitomo N,  
PI Okuda M, Saeki K;

AA  
DR WPI; 2002-437518/47.

PT New modified alkaline proteases useful in detergent compositions -

PS Claim 5; Page 15-16; 25pp; English.

This invention describes novel *Bacillus* sp. alkaline proteases useful in detergent compositions, especially in laundry, bleaching or automatic dishwasher detergents. The novel proteases have an increased detergency (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This sequence represents a fragment of the alkaline protease 1a from *Bacillus* sp strain Y-(FERM BP-1029) described in the method of the invention.

Sequence 433 AA;

Query Match	62.68;	Score 1941.5;	DB 23;	Length 433;
Best Local Similarity	86.28;	Pred. NO. 3.8e-147;		
Matches 374;	Conservative 21;	Mismatches 38;	Indels 1;	Gaps 1

QY 207 NDVARGIVKADVAQSSYGLYGQGI VAVADTGLDTGRNDSSMHEAFRGKITALYALGRTN 266

Db 1 NDVARGIVKADVAQNNYGLYGQGVVAVADTGLDTGRNDSSMHEAFRGKITALYALGRTN 60

26 / N A N D I N G H I T H V A G S V L G N G X T I N K G M A P Q A N L V F O S I M D S X G G L G L P S N L O T L F S Q A X S 326

Db	61	MSDSDNGTGTVAGSVLIG-ALNKGMAPoANLlFQSIINDSSGGGLGpLSuNTLTFSoAMN	119
QY	327	AGARlHTMSGMAALNANAYlTTDSRANVDYAKNMNTlLFAAENEXpNGGTSASpGAKKAl	386
Db	120	AGARlHTMSGMApPNAVNAIYANSROYDEIVANNMTlLFAAENEGpNGTSASpGAKKAl	179
QY	387	TYGATENLRpPFGSGYADNTlNHVAQPSNKGCTPKDGRlKRPVMApPTXlLSARSLSLAPDSSF	446
Db	180	TGATENLRpPFGSIdNNNHIAQPSNKGATRGRlKRPVNApPTFlLSARSLSLAPDSSF	239
QY	447	MANHDSKAAAYMGGSMAPIYVAGNVAOLRAPHVKNNGTIPKPSLSLKALALlAGAAAGXG	506
Db	240	MANYSKAAAYMGGSMAPIYVAGNVAOLRHPFlKNNGTlPPPSLKALKALlAGATpVGGY	299
QY	507	PNGNGGMRVTLKLSLNLVAYVNESSXlSTSoKATXlXpATlAGGRPLLSlVMSASpASTTA	566
Db	300	PNGDGMGMRVTLKLSLNLVAYVNEATALATGQKATYSfOAGQGRPLKlSLVMDADAGSTTA	358
QY	567	SYTLVNLDDLTlTAPNGXTXVYNDPXXXNNMDGNNVENVFlNXpQSGTYTlEoAYN	628
Db	360	SYTLVNLDDLTlTAPNGQKXYVNDSPYDNNMDGNNVENVFlNAQSGTYTlEoAYN	419
QY	627	VPVpGQXpSLAlVH 640	
Db	420	VPSGQRSLAlVH 433	

Search completed: April 1, 2003, 12:38:19  
Job time : 37.0274 secs

Job time : 37.0274 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: April 1, 2003, 12:37:09 ; Search time 14.0109 Seconds  
(without alignments)  
1343.997 Million cell updates/sec

Title: US-09-920-954-2

Sequence: 1 MRKKKVFSLVSAALISF.....EQQAVNPVGPQKSLAIIV 640

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2.6/plodata/1/laa/5A.COMB.pep:\*  
2: /cgn2.6/plodata/1/laa/5B.COMB.pep:\*  
3: /cgn2.6/plodata/1/laa/5A.COMB.pep:\*  
4: /cgn2.6/plodata/1/laa/5B.COMB.pep:\*  
5: /cgn2.6/plodata/1/laa/PCMTUS.COMB.pep:\*  
6: /cgn2.6/plodata/1/laa/backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3029	97.6	640	4	US-09-509-814A-6
2	3028	97.6	640	4	US-09-509-814A-8
3	3028	97.6	640	4	US-09-509-814A-4
4	3016	97.2	640	4	US-09-509-814A-2
5	3007	96.9	639	4	US-09-509-814A-1
6	2722.5	87.8	641	2	US-08-873-479-42
7	2423.5	78.1	635	2	US-08-873-479-43
8	1940.5	62.6	433	4	US-09-104-623A-4
9	1940.5	62.6	433	4	US-09-104-623A-4
10	451.5	14.6	659	4	US-08-894-818B-1
11	451.5	14.6	659	4	US-08-894-818B-1
12	408	13.2	654	4	US-08-894-818B-35
13	408	13.2	654	4	US-08-894-818B-35
14	398	12.8	659	4	US-08-894-818B-5
15	381	12.6	522	4	US-08-845-472-1
16	381	12.6	522	4	US-08-845-472-1
17	391	12.6	522	4	US-08-845-472-1
18	340	11.0	520	4	US-09-000-016-7
19	340	11.0	520	4	US-09-000-016-7
20	340	11.0	734	4	US-09-514-340-7
21	340	11.0	734	4	US-09-514-340-7
22	340	11.0	823	4	US-09-000-016-2
23	299.5	9.7	903	1	US-08-514-340-2
24	299.5	9.7	903	1	US-08-514-340-2
25	299.5	9.7	1398	1	US-08-750-532-1
26	299.5	9.7	1398	1	US-08-750-532-1
27	299.5	9.7	1398	4	US-08-894-818B-8

28	288.5	9.3	237	1	US-08-750-532-18	Sequence 18, Appl
29	282	9.1	418	2	US-08-873-479-44	Sequence 44, Appl
30	272	8.8	397	1	US-08-434-255-2	Sequence 2, Appl
31	272	8.8	397	1	US-08-434-255-4	Sequence 2, Appl
32	272	8.8	397	1	US-08-459-967-2	Sequence 2, Appl
33	272	8.8	397	1	US-08-459-967-4	Sequence 2, Appl
34	272	8.8	397	1	US-08-460-327-2	Sequence 2, Appl
35	272	8.8	397	1	US-08-460-327-4	Sequence 2, Appl
36	272	8.8	397	1	US-08-459-871-2	Sequence 2, Appl
37	272	8.8	397	1	US-08-459-871-4	Sequence 2, Appl
38	264	8.5	370	1	US-08-434-255-6	Sequence 6, Appl
39	264	8.5	370	1	US-08-459-967-6	Sequence 6, Appl
40	264	8.5	370	1	US-08-460-327-6	Sequence 6, Appl
41	264	8.5	370	1	US-08-459-871-6	Sequence 6, Appl
42	255	8.2	280	1	US-08-434-255-8	Sequence 8, Appl
43	255	8.2	280	1	US-08-459-967-8	Sequence 8, Appl
44	255	8.2	280	1	US-08-460-327-8	Sequence 8, Appl
45	255	8.2	280	1	US-08-459-871-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-09-509-814A-6  
Sequence 6, Application US/09509814A  
Patent No. 6376227  
GENERAL INFORMATION:  
APPLICANT: TAKIWA, MIKIO  
APPLICANT: OKUDA, MITSUYOSHI  
APPLICANT: SAKETI, KATSUHIISA  
APPLICANT: KUROTA, HIROMI  
APPLICANT: HITOMI, JUN  
APPLICANT: KACEYAMA, YASUSHI  
APPLICANT: SHIKATA, SHITSUMI  
APPLICANT: NOMURA, MASATOMI  
TITLE OF INVENTION: ALKALINE PROTEASE  
FILE REFERENCE: 0327-0832-0PCT  
CURRENT APPLICATION NUMBER: US/09/509, 814A  
PRIOR FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: PCT/JP98/04528  
PRIOR FILING DATE: 1998-10-07  
PRIOR APPLICATION NUMBER: JP 9-274570  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 6  
LENGTH: 640  
TYPE: PRT  
ORGANISM: Bacillus sp.  
US-09-509-814A-6  
Query Match 97.6% Score 3029; DB 4; Length 640;  
Best Local Similarity 93.3%; Pred No 4 3e-254;  
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

1 MRKKKVFSLVSAALISF.....EQQAVNPVGPQKSLAIIV 60  
1 MRKKKVFSLVSAALISF.....EQQAVNPVGPQKSLAIIV 60  
61 FLVESNVKLKXGLKKLETPVANNKLIHQFNGLPLEETKQXLETKAKLIDYIPYAY 120  
61 FLVESNVKLKXGLKKLETPVANNKLIHQFNGLPLEETKQXLETKAKLIDYIPYAY 120  
121 IVEYEGDVAXXXXXXLEHVESVEPYLPIYRIDPOLFTKGASXLVKAXALDTQKXKVEQLR 180  
121 IVEYEGDVAXXXXXXLEHVESVEPYLPIYRIDPOLFTKGASXLVKAXALDTQKXKVEQLR 180  
181 GIEIXQXXSXNDVYITAKPEYKVMNDVARGIVKADVAQSSYGLGQGIIVAAVDGLD 240  
181 GIEIXQXXSXNDVYITAKPEYKVMNDVARGIVKADVAQSSYGLGQGIIVAAVDGLD 240  
241 TGRNDSMHEAFRGKITALVALGRTNNANDTNGHGVAGSVLNGXTNKMAMQANLVE 300

```
Db 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGTVAAGSVLNGSTNKGMAPOANLVF 300
Oy 301 OSIMDSXGGLGSPNLQTLFSSQASAGARIHTNSMGAANGAYTTDSRNVDDYRKNDM 360
Db 301 OSIMDSXGGLGSPNLQTLFSSQASAGARIHTNSMGAANGAYTTDSRNVDDYRKNDM 360
Oy 361 TILFAAGNEXPNGGTISAPGAKNAITVGATENLRPSFGSYADNINHVAQFSSRPPTDGG 420
Db 361 TILFAAGNEXPNGGTISAPGAKNAITVGATENLRPSFGSYADNINHVAQFSSRPPTDGG 420
Oy 421 RIKPVMAPGRTILSARSLAPDSSFFMANHDSKYVMGTSNATPIVAGNVAQLREHFVK 480
Db 421 RIKPVMAPGRTILSARSLAPDSSFFMANHDSKYVMGTSNATPIVAGNVAQLREHFVK 480
Oy 481 NRGITPKPSLLKALIAAGADIGLGPNGNOGMRVTLDKSLNAVYVNESSLSTSQKAT 540
Db 481 NRGITPKPSLLKALIAAGADIGLGPNGNOGMRVTLDKSLNAVYVNESSLSTSQKAT 540
Oy 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVLTPAPNGTXYVGNDFXPXXNMD 600
Db 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVLTPAPNGTXYVGNDFXPXXNMD 600
Oy 601 GRNNVENFVFNKPOSGTTEIYQAVNPVPGQFSLATVN 640
Db 601 GRNNVENFVFNKPOSGTTEIYQAVNPVPGQFSLATVN 640
```

## RESULT 2

US-09-509-814A-8  
Sequence 8, Application US/09509814A

Patent No. 6376227

GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO

APPLICANT: OKUDA, MITSUYOSHI

APPLICANT: SAEKI, KATSUHIISA

APPLICANT: KUBOTA, HIROMI

APPLICANT: HITOMI, JUN

APPLICANT: KAGEYAMA, YASUSHI

APPLICANT: SHIKATA, SHITSUM

APPLICANT: NOMURA, MASAFUMI

TITLE OF INVENTION: ALKALINE PROTEASE

FILE REFERENCE: 0327-0832-0PCT

CURRENT APPLICATION NUMBER: US/09/509, 814A

CURRENT FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: PCT/JP98/04528

PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: JP 9-274570

PRIOR FILING DATE: 1997-06-08

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Patent version 3.0

SEQ ID NO 8

LENGTH: 640

TYPE: PRT

ORGANISM: Bacillus sp.

US-09-509-814A-8

Query Match 97.6%; Score 3028; DB 4; Length 640;  
Best Local Similarity 93.3%; Pred. No. 5.2e-254;

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Oy 1 MKKKKVFSLVSAAILSTVALNPSAGAXRFDLDFKGIQTTTDXGFSKQXQTGAAG 60
Db 1 MKKKKVFSLVSAAILSTVALNPSAGAXRFDLDFKGIQTTTDXGFSKQXQTGAAG 60
Oy 61 FLVESENKLLKGLKLETPANNKLIHXOFPNGPILEETKQXLEXTGAKILDIYIPDAY 120
Db 61 FLVESENKLLKGLKLETPANNKLIHXOFPNGPILEETKQXLEXTGAKILDIYIPDAY 120
Oy 121 IVEYGDVXSXXXIIEHVESVEPYLPXYXIDPOLFTKGASXYVRKXALDXTQXNKEVQLR 180
Db 121 IVEYGDVKSATSTIEHVESVEPYLPYRIDPOLFTKGASVSELVAVALDTQKXNKEVQLR 180
```

```
Oy 181 GIEIXAQQXXSNDVXYITAKPEKYKMYNDVARGIKADVAQSSYGLYGCGQIVAAVADTGLD 240
Db 181 GIEIXAQQXXSNDVXYITAKPEKYKMYNDVARGIKADVAQSSYGLYGCGQIVAAVADTGLD 240
Oy 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGTVAAGSVLNGSTNKGMAPOANLVF 300
Db 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGTVAAGSVLNGSTNKGMAPOANLVF 300
Oy 301 OSIMDSXGGLGSPNLQTLFSSQASAGARIHTNSMGAANGAYTTDSRNVDDYRKNDM 360
Db 301 OSIMDSXGGLGSPNLQTLFSSQASAGARIHTNSMGAANGAYTTDSRNVDDYRKNDM 360
Oy 361 TILFAAGNEXPNGGTISAPGAKNAITVGATENLRPSFGSYADNINHVAQFSSRPPTDGG 420
Db 361 TILFAAGNEXPNGGTISAPGAKNAITVGATENLRPSFGSYADNINHVAQFSSRPPTDGG 420
Oy 421 RIKPVMAPGRTILSARSLAPDSSFFMANHDSKYVMGTSNATPIVAGNVAQLREHFVK 480
Db 421 RIKPVMAPGRTILSARSLAPDSSFFMANHDSKYVMGTSNATPIVAGNVAQLREHFVK 480
Oy 481 NRGITPKPSLLKALIAAGADIGLGPNGNOGMRVTLDKSLNAVYVNESSLSTSQKAT 540
Db 481 NRGITPKPSLLKALIAAGADIGLGPNGNOGMRVTLDKSLNAVYVNESSLSTSQKAT 540
Oy 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVLTPAPNGTXYVGNDFXPXXNMD 600
Db 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVLTPAPNGTXYVGNDFXPXXNMD 600
Oy 601 GRNNVENFVFNKPOSGTTEIYQAVNPVPGQFSLATVN 640
Db 601 GRNNVENFVFNKPOSGTTEIYQAVNPVPGQFSLATVN 640
```

## RESULT 3

US-09-509-814A-4

Sequence 4, Application US/09509814A

Patent No. 6376227

GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO

APPLICANT: OKUDA, MITSUYOSHI

APPLICANT: SAEKI, KATSUHIISA

APPLICANT: KUBOTA, HIROMI

APPLICANT: HITOMI, JUN

APPLICANT: KAGEYAMA, YASUSHI

APPLICANT: SHIKATA, SHITSUM

TITLE OF INVENTION: ALKALINE PROTEASE

FILE REFERENCE: 0327-0832-0PCT

CURRENT APPLICATION NUMBER: US/09/509, 814A

CURRENT FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: PCT/JP98/04528

PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: JP 9-274570

PRIOR FILING DATE: 1997-06-08

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Patent version 3.0

SEQ ID NO 4

LENGTH: 639

TYPE: PRT

ORGANISM: Bacillus sp.

US-09-509-814A-4

Query Match 97.4%; Score 3020; DB 4; Length 639;  
Best Local Similarity 93.4%; Pred. No. 2.6e-253;

Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Oy 4 KKKVFLVLSAAAILSTVALNPSAGAXRFDLDFKGIQTTTDXGFSKQXQTGAAGFLV 63
Db 3 KKKVFLVLSAAAILSTVALNPSAGAXRFDLDFKGIQTTTDXGFSKQXQTGAAGFLV 62
Oy 64 ESENVKLLKGLKLETPANNKLIHXOFPNGPILEETKQXLEXTGAKILDIYIPDAYIVE 123
Db 63 ESENVKLLKGLKLETPANNKLIHXOFPNGPILEETKQXLEXTGAKILDIYIPDAYIVE 122
```



Qy	124	YEDDYKXXXXXXEHESEVEPLPYXXIDPOLFTKGSXLYAAALDTRKXNKEVQJRGIE	183	OTHER INFORMATION: Xaa	is any amino acid
Db	123	YEGDQVQSKVRSIEHVESVEPLPYKXIDPOLFTKGSITLVKALDTRKXNKEVQJRGIE	182	NAME/KEY: misc_feature	LOCATION: (33)..(33)
Qy	184	XIXQXXXNDVYYITAKREYVYVMDVARGVYADVQSSYGLYQGGQIVAVADTGLDTR	243	NAME/KEY: misc_feature	LOCATION: (47)..(47)
Db	183	EIMQVYASNDVHYITAKREYVYVMDVARGVYADVQSSYGLYQGGQIVAVADTGLDTR	242	OTHER INFORMATION: Xaa	is any amino acid
Qy	244	NDSSEHAEFRKCTITVYALGFTNANDRNGHGTTHVAGSYLNGKTKNGAPQANLVEOSI	303	NAME/KEY: misc_feature	LOCATION: (48)..(48)
Db	243	NDSSEHAEFRKCTITVYALGFTNANDRNGHGTTHVAGSYLNGKTKNGAPQANLVEOSI	302	OTHER INFORMATION: Xaa	is any amino acid
Qy	304	MDSXGGLGGLPSNLOTLEFSQAXSAGARIHTNSGAAVNGAYTTTSSRVNDVYVRKNDMTLL	363	NAME/KEY: misc_feature	LOCATION: (54)..(54)
Db	303	MDSXGGLGGLPSNLOTLEFSQAFSGARIRHTNSGAAVNGAYTTTSSRVNDVYVRKNDMTLL	362	OTHER INFORMATION: Xaa	is any amino acid
Qy	364	FAAGNEXPNGGCTISAPGTAKNATIVGATENLRPSFGSYADNINNHVAFSSRGPTKGRGRIK	423	NAME/KEY: misc_feature	LOCATION: (71)..(71)
Db	363	FAAGNEXPNGGCTISAPGTAKNATIVGATENLRPSFGSYADNINNHVAFSSRGPTKGRGRIK	422	OTHER INFORMATION: Xaa	is any amino acid
Qy	424	PDVNAAGCTIISARSLSLAPDSSFPANHDSKYAVMGSTMAPIYAGVAVDLRHPFKNG	483	NAME/KEY: misc_feature	LOCATION: (75)..(75)
Db	423	PDVNAAGCTIISARSLSLAPDSSFPANHDSKYAVMGSTMAPIYAGVAVDLRHPFKNG	482	OTHER INFORMATION: Xaa	is any amino acid
Qy	484	ITPKPSLLKALILAGADXLGLTPNGNGKGRVTLDDKSLNVAYNSSXSLSTQKATYFE	543	NAME/KEY: misc_feature	LOCATION: (90)..(90)
Db	483	ITPKPSLLKALILAGADXLGLTPNGNGKGRVTLDDKSLNVAYNSSXSLSTQKATYFE	542	OTHER INFORMATION: Xaa	is any amino acid
Qy	544	TATAGPPLKSLIWSMDAPASTTASVTLVNDLVLITAPNGTYVYVGNDFXXXNNDGRN	603	NAME/KEY: misc_feature	LOCATION: (103)..(103)
Db	543	TATAGPPLKSLIWSMDAPASTTASVTLVNDLVLITAPNGTYVYVGNDFXXXNNDGRN	602	OTHER INFORMATION: Xaa	is any amino acid
Qy	604	NVENVEINXNPOSGTYTTEVQAYVNPVGPQXFSIAIVN	640	NAME/KEY: misc_feature	LOCATION: (106)..(106)
Db	603	NVENVEINXNPOSGTYTTEVQAYVNPVGPQXFSIAIVN	639	OTHER INFORMATION: Xaa	is any amino acid
RESULT 4					
US-09-509-814A-2					
Sequence 2, Application US/09509814A					
Patent No. 6376227					
GENERAL INFORMATION:					
APPLICANT: TAKAIWA, MIKIO					
APPLICANT: OKUDA, MITSUYOSHI					
APPLICANT: SAEKI, KATSUHIISA					
APPLICANT: KUBOTA, HIROHI					
APPLICANT: HITOMI, JUN					
APPLICANT: KAGEYAMA, YASUSHI					
APPLICANT: SHIKATA, SHITSUM					
APPLICANT: NOMURA, MASAYUMI					
TITLE OF INVENTION: ALKALINE PROTEASE					
FILE REFERENCE: 0347-0833-00CT					
CURRENT APPLICATION NUMBER: US/09/509, 814A					
CURRENT FILING DATE: 2000-04-06					
PRIOR APPLICATION NUMBER: PCT/JP98/04528					
PRIOR FILING DATE: 1998-10-07					
PRIOR APPLICATION NUMBER: JP 9-274570					
PRIOR FILING DATE: 1997-06-08					
NUMBER OF SEQ ID NOS: 24					
SOFTWARE: PatentIn version 3.0					
SEQ ID NO 2					
LENGTH: 640					
TYPE: PRT					
ORGANISM: Bacillus sp.					
FEATURE:					
NAME/KEY: misc_feature					
LOCATION: (3)..(3)					
OTHER INFORMATION: Xaa is any amino acid					
NAME/KEY: misc_feature					
LOCATION: (24)..(24)					
OTHER INFORMATION: Xaa is any amino acid					
NAME/KEY: misc_feature					
LOCATION: (195)..(195)					

```

: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (287)..(287)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (307)..(307)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (325)..(325)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (370)..(370)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (432)..(432)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (502)..(502)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (532)..(532)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (542)..(542)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (585)..(585)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (592)..(592)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (593)..(593)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (595)..(595)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (596)..(596)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (597)..(597)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (612)..(612)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (633)..(633)
: OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2

```

```

Query Match          97.2%; Score 3016; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 5,7e-253;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MRKKKVELSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTITDXXGFSKXOTGAA 60
DB 1 MRKKKVELSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTITDXXGFSKXOTGAA 60
QY 61 FLVSEENVKLKGLKKLETPANNKLIHXQFNGPILEETKXOLEXTGAKIIDYIPDAY 120
DB 61 FLVSEENVKLKGLKKLETPANNKLIHXQFNGPILEETKXOLEXTGAKIIDYIPDAY 120
QY 121 IVEYGDVXSXXXIIEHVESVEPYLPXYXIDPOLFTKGSXIVKXALDTCXNKNEVOLR 180
DB 121 IVEYGDVXSXXXIIEHVESVEPYLPXYXIDPOLFTKGSXIVKXALDTCXNKNEVOLR 180
QY 181 GIEIXAQXXXSNDVYITAKPEYKVMNDVARGIVKADVAQSSYGLYGOGIYAADVDTGLD 240
DB 181 GIEIXAQXXXSNDVYITAKPEYKVMNDVARGIVKADVAQSSYGLYGOGIYAADVDTGLD 240
QY 241 TGRNDSMHEAFRGKITALVALGRTNANDTNGHGTAVAGSVLNGXTNKGMAPQANLVF 300
DB 241 TGRNDSMHEAFRGKITALVALGRTNANDTNGHGTAVAGSVLNGXTNKGMAPQANLVF 300

```

```

DB 241 TGRNDSMHEAFRGKITALVALGRTNANDTNGHGTAVAGSVLNGXTNKGMAPQANLVF 300
QY 301 OSIMDSXGGLGCLPSNLOTLEFSQASAGARITHNSMGAAYVNTDTSRNVDDYVRKNDM 360
DB 301 OSIMDSXGGLGCLPSNLOTLEFSQASAGARITHNSMGAAYVNTDTSRNVDDYVRKNDM 360
QY 361 TLLEAAGNEXPNGGTISAPGTRAKNAITVGARENLPRSEGSADNINNHVAOFSSRCPTKD 420
DB 361 TLLEAAGNEXPNGGTISAPGTRAKNAITVGARENLPRSEGSADNINNHVAOFSSRCPTKD 420
QY 421 RIKPDVMAFGTYIISARSLSLAPDSSFMANHDSKTAIVMGTSMAITIVAGNVAQLREHFVK 480
DB 421 RIKPDVMAFGTYIISARSLSLAPDSSFMANHDSKTAIVMGTSMAITIVAGNVAQLREHFVK 480
QY 481 NRGITPKPSLLKALIAAGADXLGPGNCGMGRVTLTDLKSLNVAVNESSXLSTSORAT 540
DB 481 NRGITPKPSLLKALIAAGADXLGPGNCGMGRVTLTDLKSLNVAVNESSXLSTSORAT 540
QY 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVTTPANGTYVGNDEFPXXKXMD 600
DB 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVTTPANGTYVGNDEFPXXKXMD 600
QY 601 GRNNVENVFINKPQSGTYTIEVOAYNVPGQXPSLATVN 640
DB 601 GRNNVENVFINKPQSGTYTIEVOAYNVPGQXPSLATVN 640

```

```

RESULT 5
US-09-509-814A-1
: Sequence 1, Application US/09509814A
: Patent No. 6376227
: GENERAL INFORMATION:
: APPLICANT: TAKAIWA, MIKIO
: APPLICANT: OKUDA, MITSUYOSHI
: APPLICANT: SAKKI, KATSUHIISA
: APPLICANT: KUBOTA, HIROMI
: APPLICANT: HITOMI, JUN
: APPLICANT: KAGEYAMA, YASUSHI
: APPLICANT: SHIKATA, SHITSUW
: APPLICANT: NOMURA, MASAFUMI
: TITLE OF INVENTION: ALKALINE PROTEASE
: FILE REFERENCE: 0327-0832-0PCT
: CURRENT APPLICATION NUMBER: US/09/509, 814A
: PRIOR FILING DATE: 2000-04-06
: PRIOR APPLICATION NUMBER: PCT/JP98/04528
: PRIOR FILING DATE: 1998-10-07
: PRIOR APPLICATION NUMBER: JP 9-274570
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1
: LENGTH: 639
: TYPE: PRT
: ORGANISM: Bacillus sp.
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (23)..(23)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (29)..(29)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (32)..(32)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (46)..(46)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (47)..(47)
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (53)..(53)
: OTHER INFORMATION: Xaa is any amino acid

```

```
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (185)..(185)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
```

```
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match          96.9%  Score 3007;  DB 4;  Length 639;
Best Local Similarity 100.0%;  Pred. No. 3,4e-252;
Matches 637;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY  4  KKKVFLSVLSAAAILSTVALXNPSGAXRFXFDLDFKGIOTTTDXGFSKXOTGAALFLV 63
DB  3  KKKVFLSVLSAAAILSTVALXNPSGAXRFXFDLDFKGIOTTTDXGFSKXOTGAALFLV 62
QY  64  ESENVRKLXGKLKKLETVPANNNKLHIXQFNGLPIEETKQXLEXTGAKILDYIPDYAYIVE 123
DB  63  ESENVRKLXGKLKKLETVPANNNKLHIXQFNGLPIEETKQXLEXTGAKILDYIPDYAYIVE 122
QY  124  YEGDVYXXXXXIEHVESVEPYLPYXXIDPOLFTKGASXLVKAXALDTQXKKEVQLMGIE 183
DB  123  YEGDVYXXXXXIEHVESVEPYLPYXXIDPOLFTKGASXLVKAXALDTQXKKEVQLMGIE 182
QY  184  XIAQXXXXNDVYITRAKREYVMDVARGIVKADVQSSYGLYGGOQIVAAVADTGLDTGR 243
DB  183  XIAQXXXXNDVYITRAKREYVMDVARGIVKADVQSSYGLYGGOQIVAAVADTGLDTGR 242
QY  244  NDSSMHEAFRGKITALVALGRTNANDTNGHGVAGSVLGNXTNKMADQANLVFOST 303
DB  243  NDSSMHEAFRGKITALVALGRTNANDTNGHGVAGSVLGNXTNKMADQANLVFOST 302
QY  304  MDSXGILGGLPSNLQTLFSSQAXSAGARIHTNSGAAVNCAVYTTDSRNVDDVYRRKNDMTIL 363
DB  303  MDSXGILGGLPSNLQTLFSSQAXSAGARIHTNSGAAVNCAVYTTDSRNVDDVYRRKNDMTIL 362
QY  364  FAAGNEXXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNTNHNVAOPSSGPGTKDRIK 423
DB  363  FAAGNEXXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNTNHNVAOPSSGPGTKDRIK 422
QY  424  PDVMARCTYIILRSASLAPDSSFPANHDSKYAAMGCTSMATPIVAGNVAOLREHFVNRG 483
DB  423  PDVMARCTYIILRSASLAPDSSFPANHDSKYAAMGCTSMATPIVAGNVAOLREHFVNRG 482
```

QY 484 ITPKPSLLKALIALAGADGGLGYPNGOGWGRVTLDKSLNAVYNSSXLSTSQKATYXF 543  
DB 483 ITPKPSLLKALIALAGADGGLGYPNGOGWGRVTLDKSLNAVYNSSXLSTSQKATYXF 542  
QY 544 TATACKPLKISLWSDAPASTASTATLVNDLDTTAPNGTYVYVNDPFXPPXXWMDGRN 603  
DB 543 TATACKPLKISLWSDAPASTASTATLVNDLDTTAPNGTYVYVNDPFXPPXXWMDGRN 602  
QY 604 NVENFINKPQSGTYTIEVOAYNPVGPQXFSALIVN 640  
DB 603 NVENFINKPQSGTYTIEVOAYNPVGPQXFSALIVN 639

## RESULT 6

US-08-873-479-42  
Sequence 42, Application US/08873479  
Patent No. 5891701  
GENERAL INFORMATION:  
APPLICANT: Sloma, Alan  
APPLICANT: Lynne, Christlanson  
TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/873.479  
FILING DATE: 12-JUN-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Agilis, Cheryl H  
REGISTRATION NUMBER: 34,086  
REFERENCE/DOCKET NUMBER: 5251.000-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
TELEX:  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 641 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-873-479-42

Query Match 87.8%; Score 2722.5; DB 2; Length 641;  
Best Local Similarity 82.9%; Pred. No. 1.5e-227;  
Matches 532; Conservative 40; Mismatches 67; Indels 3; Gaps 2;

QY 1 MRKX--KKVFLSVLSAAAILSTVALXNPSSAGXARXFDLDFKGIQTFTTXXGFSKQXQGA 58  
DB 1 MRKSKSRVFLSVLSVALLSVALSPSTIGANNPELDFKGIETLTLEKATKQKGRK 60  
QY 59 AAFVLESERVKILKGLKLETVPANNNKLUHQFNGPILEETKQKLEXTGAKILDIYIDY 118  
DB 61 ASFLVNSERVKIPKSIQKLEVPADNKLIVQFDGPILLETQLOLETKGAKILDIYIDY 120  
QY 119 AYIYEDGVYKXXXXXIEHVESVEPYLDPXYXIDPOLFTKGSXLKAKALDTKQXNKQ 178  
DB 121 AYIYEDGVYKAVTNAIHLSEVEPYLDPYKIDPOLFSKGSSELVEYVALDKKQSK 180

QY 179 LRGIEIAOXXXXNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQOIVAVADTG 238  
DB 181 LRGLEQIAQVATNNNDLVYTPKPEYEVLLNDVARGIVKADVAQNNNGGLYGQOIVAVADTG 240  
QY 229 LDTGRNDSMMHEAFKGTITALLYALGRTNNANDTNHGHGVAGSVLGNKXTKKGAPOANL 298  
DB 241 LDTGRNDSMMHEAFKGTITALLYALGRTNNANDTNHGHGVAGSVLGN-ATNKGMAPOANL 299  
QY 229 VFQSTMDSGXGGLGJPSNLQTLFISOAXSAGARIHNSNGAAVYTTDSRRVDDYVRKN 358  
DB 300 VFQSTMDSGGGLGJLPANLQTLFISOAYSGARIHNSNGAPVNGAYTTDSRRVDDYVRKN 359  
QY 359 DMTILFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADINHNVAQFSSRGPTK 418  
DB 360 DMTILFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADINHNVAQFSSRGPTK 419  
QY 419 DGRIPDYMAPGTXYLSARSSLAPDSSFWANHDSKYATMGTSMTPIYAGVAVOLREHF 478  
DB 420 DGRIPDYMAPGTXYLSARSSLAPDSSFWANHDSKYATMGTSMTPIYAGVAVOLREHF 479  
QY 479 VKNRGITPKPSLLKALIALAGADGGLGYPNGOGWGRVTLDKSLNAVYNSSXLSTSQK 538  
DB 480 VKNRGITPKPSLLKALIALAGADGGLGYPNGOGWGRVTLDKSLNAVYNSSXLSTSQK 539  
QY 539 ATYXFTATAGKPLKISLWSDAPASTASTATLVNDLDTTAPNGTYVYVNDPFXPPXXN 598  
DB 540 ATYXFTATAGKPLKISLWSDAPASTASTATLVNDLDTTAPNGTYVYVNDPFXPPXXN 599  
QY 599 WGRNNVENVFINKPQSGTYTIEVOAYNPVGPQXFSALIVN 640  
DB 600 WGRNNVENVFINKPQSGTYTIEVOAYNPVGPQXFSALIVN 641

## RESULT 7

US-08-873-479-43  
Sequence 43, Application US/08873479  
Patent No. 5891701  
GENERAL INFORMATION:  
APPLICANT: Sloma, Alan  
APPLICANT: Lynne, Christlanson  
TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/873.479  
FILING DATE: 12-JUN-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Agilis, Cheryl H  
REGISTRATION NUMBER: 34,086  
REFERENCE/DOCKET NUMBER: 5251.000-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
TELEX:  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 635 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-873-479-43

## Query Match

78.1%; Score 2423.5; DB 2; Length 635;  
Best Local Similarity 73.1%; Pred. No. 1.1e-201;

Matches 468; Conservative 68; Mismatches 99; Indels 5; Gaps 4;

```
QY 1 MRKKKVFSLVSAALISTVALKNSGAXRFXDDEFKGIOTTTDXGFSKXOTGAAA 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MKGKKRVLSVSAALIASVWSSPTSGA--DFQVFNQGVK-SLENSALVPISSGBAS 57
QY 61 FVSESEVWKLXGKLLKLETPANNKLIHQFNGPILTEERKXOLETGAKLIDTPYAY 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 58 FVDTENINPFGIOKLEAVQKXDELITYQFPGISEERKGLSLGSLIDYPPYAF 117
QY 121 IVEEDGVYKXXXIIEHVESEPTLPYXIDPOLTFKGSALYKXALDTPQXKEVOLR 180
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 118 IVQYSG-ATKRNSTLSHVENQVPELTKIDPELTKIGMSQLQVAILMTKHNKMKFT 176
QY 181 GIEKIXQXXXNDVYITRAKEPKYKNDVARGIYKADVAQSSYGLYGQGIYAADTGLD 240
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 177 GLDEIVQYANDVLTISPKEPEYELMDVARGIYKADVAQNNNGLYGQGIYAADTGLD 236
QY 241 TGRNDSMHEAFRGKITALYALGRTNNANDTNGHGHVAGSVLNGXTNKGAPQANLYF 300
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 237 TGRNDSMHEAFRGKITALYALGRTNNASDPNGHGHVAGSVLGN-ALNKGMAPQANLYF 295
QY 301 QSIDNXGGLGCLPSNLQTLFSGAXSAGARIHTNSGAAVNGAYTTDSRNVDDYRKNDM 360
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 296 QSIDNDSGGIGLGLPSNLMTLFSQAMNAGARIHTNSMGAPVNGAYTANSRQDEYVRNNDM 355
QY 361 TIEFAAGNEXPGNGTISAPGTAKNATVYGAENLRPSFGSYADNINHVAFSSRGPTKDG 420
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 356 TIEFAAGNEXPGNGTISAPGTAKNATVYGAENLRPSFGSIDNPNHIAQSSRGATRDG 415
QY 421 RIRPDVMAAGTYIISABSLAPDSSFPANHDSKAYAWGCSMATPIYAGNVAQLREHFVK 480
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 416 RIRPDVMAAGTYIISABSLAPDSSFPANHDSKAYAWGCSMATPIYAGNVAQLREHFVK 475
QY 481 NRGITPEPSSLKALIALAGADXLGYTPNGNQGWRVTLQKSLNVAVYESSXLTSSOKAT 540
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 476 NRGITPEPSSLKALIALAGADXLGYTPNGNQGWRVTLQKSLNVAVYESSXLTSSOKAT 535
QY 541 YXETATAGRPKLSLWSDAPASTASYTLVNDLVTAPNGTYXVGNDFXPXKXNMD 600
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 536 YXFOAQAGRPKLSLWSDAPASTASYTLVNDLVTAPNGTYXVGNDFXPXKXNMD 595
QY 601 GRNNVEVFVIXXPOSGTYTIEVOAVNVVGPQXFSIAIYN 640
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 596 GRNNVEVFVIXXPOSGTYTIEVOAVNVVGPQXFSIAIYN 635
```

## RESULT 8

US-09-104-623A-4  
Sequence 4, Application US/09104623A  
Patent No. 6303752

## GENERAL INFORMATION:

APPLICANT: Olsen, Arne Agerlin  
APPLICANT: Fatum, Tine Muxoll  
APPLICANT: Deussen, Helmut-Josef  
APPLICANT: Roggen, Erwin Ludo  
TITLE OF INVENTION: A Modified Polypeptide  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 63037520 No. 6303752disk of No. 6303752th America, Inc.  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS

SOFTWARE: FASTED for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/104,623A

FILING DATE: 25-JUN-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rozek, Carol

REGISTRATION NUMBER: 36,993

REFERENCE/DOCKET NUMBER: 5256,200-US

TELEPHONE: 212-867-0123

TELEFAX: 212-867-9655

INFORMATION FOR SEQ ID NO:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

STRAIN: Bacillus sp. Y

US-09-104-623A-4

## Query Match

62.6%; Score 1940.5; DB 4; Length 433;

Best Local Similarity 86.2%; Pred. No. 4.3e-160;

Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

```
QY 207 NDVARGIVKADVAQSSYGLYGQGIYVAVADTGTGDRNDSMHEAFRGKITALYALGRTN 266
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 NDVARGIVKADVAQNNNGLYGQGIYVAVADTGTGDRNDSMHEAFRGKITALYALGRTN 60
QY 267 NANDTNGHGHVAGSVLNGXTNKGAPQANLYFOSIMSGXGGLGCLPSNLQTLFSGAXS 326
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NASDPNGHGHVAGSVLGN-ALNKGMAPQANLYFOSIMSGXGGLGCLPSNLQTLFSGAXS 119
QY 327 AGARHTNSGAAVNGAYTTDSRNVDDYRKNDMTILFPAAGNEXPGNGTISAPGTAKNAT 386
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 AGARHTNSGAAVNGAYTANSRQDEYVRNNDMTILFPAAGNEXPGNGTISAPGTAKNAT 179
QY 387 TYGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAAGTYIISABSLAPDSSF 446
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 TYGATENLRPSFGSIDNPNHIAQSSRGATRDGRIKPDVMAAGTYIISABSLAPDSSF 239
QY 447 WANHDSKVAYMGCSMATPIYAGNVAQLREHFVKNRGITPKPSLIKALIALGADVLGY 506
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 WANHDSKVAYMGCSMATPIYAGNVAQLREHFVKNRGITPKPSLIKALIALGADVLGY 299
QY 507 PNCNOGWRVTLQKSLNVAVYESSXLTSSOKATYXETATAGRPKLSLWSDAPASTTA 566
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 PNCNOGWRVTLQKSLNVAVYESSXLTSSOKATYXETATAGRPKLSLWSDAPASTTA 359
QY 567 SVTLVNDLVTAPNGTYXVGNDFXPXKXNMDGRNNVEVFVIXXPOSGTYTIEVOAVN 626
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 SVTLVNDLVTAPNGTYXVGNDFXPXKXNMDGRNNVEVFVIXXPOSGTYTIEVOAVN 419
QY 627 VPVGQXFSIAIYN 640
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 420 VPVGQXFSIAIYN 433
```

## RESULT 9

US-09-019-532-4  
Sequence 4, Application US/09019532B  
Patent No. 6416756

## GENERAL INFORMATION:

APPLICANT: Olsen, Arne Agerlin  
APPLICANT: Prent, Annette  
TITLE OF INVENTION: A Modified Enzyme for Skin Care  
FILE REFERENCE: 4922,204-US  
CURRENT APPLICATION NUMBER: US/09/019,532B  
CURRENT FILING DATE: 1998-02-05  
EARLIER APPLICATION NUMBER: 0038/97  
EARLIER FILING DATE: 1997-01-10

EARLIER APPLICATION NUMBER: 0754/97  
EARLIER FILING DATE: 1997-06-25  
EARLIER APPLICATION NUMBER: 60/051,381  
EARLIER FILING DATE: 1997-07-07  
EARLIER APPLICATION NUMBER: PCT/DK98/00015  
EARLIER FILING DATE: 1998-01-12  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 433  
TYPE: PRT  
ORGANISM: Bacillus sp.  
US-09-019-532-4

Query Match 62.6%; Score 1940.5; DB 4; Length 433;  
Best Local Similarity 86.2%; Pred. No. 4,3e-160;  
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 207 NDVARGIKADVAAOSSYGLGOGIIVAVADTGLDTRNDSSMHEAFRGKITALVALGRTN 266  
DB 1 NDVARGIKADVAAONNGLYGOGIIVAVADTGLDTRNDSSMHEAFRGKITALVALGRTN 60  
QY 267 NANTNGHGTIVAGSVLGNKXTNGMAPQANLVFOSIMDSXGUGLPSNLQTLFSQAXS 326  
DB 61 NASDPNGHGTIVAGSVLGN-ALNKGMAPQANLVFOSIMDSXGUGLPSNLQTLFSQAMN 119  
QY 327 AGARIHTSMGCAVANGATYTRNDVYVRKNDMTILFAAGNEXNGGTISAPGTAKNAI 386  
DB 120 AGARIHTSMGAPVNGATYTRNDVYVRKNDMTILFAAGNEXNGGTISAPGTAKNAI 179  
QY 387 TVGATENLRPSFGSYADININHAOFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSSF 446  
DB 180 TVGATENLRPSFGSYADININHAOFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSSF 239  
QY 447 MANDSKYAVYNGTSMATPIYAGVNAOLREHFNKRGITPRSLIKALTINGADHXGIGY 506  
DB 240 MANDSKYAVYNGTSMATPIYAGVNAOLREHFNKRGITPRSLIKALTINGADHXGIGY 299  
QY 507 PNGNGMGWRLTDLKSLNVAAYNESSXISTSQKATYXTATAGKPLKISLWSDAPASTTA 566  
DB 300 PNGNGMGWRLTDLKSLNVAAYNESSXISTSQKATYXTATAGKPLKISLWSDAPASTTA 359  
QY 567 SVTLVNDLDTITAPNGXYVGNDEXXPXXNNMGRNNEVNFITNPGSGTYTIEVQAYN 626  
DB 360 SVTLVNDLDTITAPNGXYVGNDEXXPXXNNMGRNNEVNFITNPGSGTYTIEVQAYN 419  
QY 627 VPGVQXESLAIYN 640  
DB 420 VPGVQXESLAIYN 433

RESULT 10  
US-08-894-818B-1  
Sequence 1, Application US/08894818B  
Patent No. 6261822  
GENERAL INFORMATION:  
APPLICANT: TAKAKURA, Hikaru  
APPLICANT: MORISHITA, Mio  
APPLICANT: YAMAMOTO, Katsuhiko  
APPLICANT: MITTA, Masanori  
APPLICANT: ASADA, Kiyozo  
APPLICANT: TSUNASAWA, Susumu  
APPLICANT: KATO, Ikunoshin  
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Broadway and Neimark  
STREET: 419 Seventh Street N.W., Ste. 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: United States of America  
ZIP: 20004  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,818B  
FILING DATE: 20-May-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03253  
FILING DATE: 07-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 323285/1995  
FILING DATE: 12-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Broadway, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: TAKAKURA-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-5197  
TELEFAX: (202) 737-3528  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 659 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-818B-1

Query Match 14.6%; Score 451.5; DB 4; Length 659;  
Best Local Similarity 25.9%; Pred. No. 6.1e-31;  
Matches 157; Conservative 92; Mismatches 207; Indels 151; Gaps 24;

QY 70 LKGIKKLELVPPANNKHIXQFNGPILEETK--OXLEXTGATIDTIDPYATIVEREGD 127  
DB 43 LTPGLFKVQAMNNQEVDTYIMGSGYDRBRAKVLRLMGAQV-----KSYKI----- 92  
QY 128 VXSXXXXIEHVESVEPYLPXYXIDPQLFTKGASKLVKAXALDTQXNKEVOLRIEXIAQ 187  
DB 93 -----IPAAVAK-----IKARDLLTAGMIDTGYFG-NRFVSGIKFIOE 130  
QY 188 XXXSNDVYITAKPEYKMNDAVARGI--VRADVAOSSGLGOGIIVAVADTGLDTRND 245  
DB 131 -----DYKQVDDATSVQIGADYVWNSLGDGSGVVAIVDTGIDAN--- 173  
QY 246 SSMHEAFRGKITALY-ALGRTNNANDTNGHSTHVASVYLGNGXTNK---GAPQANLVFQ 301  
DB 174 ---HPDLKGYIYGNWDANNGRSTPYDDGCHGTIVAGTGSVNSQYIGVAPGAKLVGV 230  
QY 302 SIM--DSXGGLGLPSNLQTLFSQAXSAGARI-----HTNSGCAVNGAYTTD 347  
DB 231 KVLGADSGSVSTIAGDVWVQNKDKYGIIVINLSLSSQSSDGTDSLQAVNNAMDA- 289  
QY 348 SRNDDVYRKNDMTILFAAGNEXNGGTISAPGTAKNAITVGATENLRPSGSAADNINH 407  
DB 290 -----GIIVCVAAAGNSGPRYTYVGSPPAASKYITVGA-----VDSNDN 327  
QY 408 VAOFSRGPTRKDGRIKPDVMAFGTYILSARSLAPDSSFMANHNSKYAVMGTSMATPIY 467  
DB 328 IASFSRGPTRADGRLKPEYVAVPGVDIAPRAS---GTSMGPIINDYITKAGTSMATPHV 384  
QY 468 AGVNAQLREHFNKNGITP---KPSLAKA-LIAGAADXLGYPNGMGWRLTDLKSL- 522  
DB 385 SGVGLLIQ---AHPSPWPRDKVKTALLETADIVAPKEIADIAY-----GAGRNVVYKAK 436  
QY 523 --NVAAYNESSXISTSQKATYXTATAGKPLKISLWSDAPASTTASVTLVNDLDTITA 580  
DB 437 YDDYAKLTFGTSVADKSGATHTFPVSGATFVATATLYMD-----TGSSDIDLFLYD 486  
QY 581 PNG-----TXYYGNDPFXXPXXNNMGRNNEVNFITNPGSGTYTIEVQAYNPGVQX 633  
DB 487 PNGNEVDISTATYIG-----FEKVGYYNPTAGTWTYKVVYSYK---GAAN 527

```

0Y 70 LKGLKMKLETFYVANNKLLHIXOFNGPBLEETK--OXLEXTAKLIIDYIDPYAVIYEBCD 127
Db 43 LTPGLFEKKVORMNMDEVDYIMFGSYGDRVAVYKILMGAQV-----KYSYKI----- 92
QY 128 VXSXXXIXIEHVESVEPEYLPYXIIDPOLFTFGASXLVKAAALDTQXNKEVOJLGEIXIAQ 187
Db 93 -----IPAAVAK-----IKAROLLILIGIMIDTGYEFG-MNPVSGIKPFIOE 130
QY 188 XXXSNDXYITTAKEPKYKVNADVARGI--VKADVAOSVGYLGQGOQIYAANADLTGTGBND 245
Db 131 -----DKVQVDDATSVSIOIADPTVHNSLGYGSGSVVAIIVDYGIDAN--- 173
QY 246 SSMHEAFRGKITALY-ALGRTNNNDNDNGHGTGHVAGSVLGNCGYXNK-----GMAPOLNLYFO 301
Db 174 ---HPDEKGVKVIWMYDVAVNGSRPTPYDQGGTGHVAGTGSVNSQYIGVAPGAKLGV 230
QY 302 SIM--DSXGGLGLPBNLOTLEFSQASAGARI-----HNSMGAANVCAATTD 347
Db 231 KYLGADGSGSVTIIINGVDVWVOKKDKYGRIVINILSLGSSQSDCTDLSLQAVNANAMA- 289
QY 348 SHNDVDYVRKMDTITFIAGAXENPGSTISAPETKMNITGATENLRNRFSGSYADINNH 407
Db 290 -----GLVYCAVGAHSGPNYITTVGSPAAASKYIITGA-----VDSMDN 322
QY 408 VAOSSSGPKKDRIRKPDVMAPTXILIASSLAADSPFMANHDSKAYACGSMATPIY 467
Db 328 IASTSSSGPFLADRLKEPEVAVPVDIITAPAS---GTSMGPIINDIYTKASGTSMAITHV 384
QY 468 AGNVAOLREHFVANKRPII--KSLILKA-LIAGAADGGLCTPNNGMGKRYTLIDKSL- 522
Db 385 SGVCGALLQ--AHSPMTPDKAKTALLERLDIYAKRELAIDAI---GAGRANVYKAIK 436
QY 523 ---NAYVNESSXJSTQAKATYXFTATYAGKRLKLSLYWSDAPASTASVYLVNDLVLITA 580
Db 437 YDDIAKLFTGVSADKSGHTRHFEVSQATRYATILKWD-----GSSDIDLVLTD 486
QY 581 PNG-----TXYGNDYEXXPKXXNMGRNANVENFIAKPOSGTYTIEVOATYVNPVGOX 633
Db 487 PGNNEVDYSTATIYG-----EKGYTNPAGTWYKVVSTK-----GAAN 527

```

[illegible]

RESULT 13  
 US-09-445-472-16  
 : Sequence 16 Application US/09445472  
 : Patent No. 6358726  
 : GENERAL INFORMATION:  
 : APPLICANT: TAKAKURA, Hikaru  
 : APPLICANT: MORISHITA, Mio  
 : APPLICANT: SHIMOJO, Tomoko  
 : APPLICANT: ASADA, Kiyozo  
 : APPLICANT: KATO, Ikunoshin  
 : TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE  
 : FILE REFERENCE: TAKAKURA-6  
 : CURRENT APPLICATION NUMBER: US/09/445,472  
 : CURRENT FILING DATE: 1999-12-06  
 : PRIOR APPLICATION NUMBER: 151969/1997  
 : PRIOR FILING DATE: 1997-06-10  
 : NUMBER OF SEO ID NOS: 33  
 : SOFTWARE: PatentIn version 3.0  
 : SEO ID NO 16  
 : LENGTH: 654  
 : TYPE: PRT  
 : ORGANISM: PYROCOCCUS FURIOSUS  
 : US-09-445-472-16

[illegible]

Qy	241	TGRDSSMHEHFRCKITATYALVLTGNTNAN-----	DTNGCHTHAGSVLNGKXTN-----	K	290
Db	171	-----ASHPDQGV-----	TGVNDDVNGSRPYDDHGHGHTHASTAACTGAASNGKY	219	
Qy	291	GMAPQAMLVFQSIM--DSXGIGLGLPSNLQTLFQSA	XSAARJHTNSWGA-----	338	
Db	220	GMAGCAKLAGIKVLGADSGSISITTINGEVA	VDNKKQYGIKVINILSGSSQSDGTAL	279	
Qy	339	--AVNGAYTDSRVVDYVYRKNDMTILFPA	GNEXEPNGGTTISAGCTAKNALTIVGATEMLP	356	
Db	280	SOAVNAAMD-----	GLVVVYVAGNSGPNKTYTTGSP	LAASKVTVGA-----	321
Qy	397	SFGSVADMINHVAQFSSNGPKKDKRIPDVA	PGTXYLXLRSSLIAPSSFWANHDSKYA	456	
Db	322	-----VDKYDVTITFSSHGCPADRLKPE	VYVAPGNMIIAARAS--GTSMQRPINDYTTA	373	
Qy	457	MGGTSMATPIYAGVNAOLREHFVKNRGITPK--	PSLKAALITAGA-----	ADXGLGY	506
Db	374	APGTSMATPIHVAIGAILLLQ-----	AHPSWPFDKVKTALIFTADIVKDPDEIAD--	IAY	424
Qy	507	PNGMGKRYTLDKSLNVAAYNESSXLS	TSDKA-----TYXFTATAGKPLKISLWS	DAP	561
Db	425	-----GAGRVAAYKRAIN--VDNYAKLV	FTGYVANKSGQTHQFYVSGASFTATLYMDNAN	477	
Qy	562	ASTASVTLVNDLDTLVITAPNG-----	TXVGNDEXPXKXNMGDNRNVENVFI	IXPO	614
Db	478	-----SDLDLTYLDPDNGNVDSYTA	YVY-----FEKGYVNP	511	
Qy	615	SGTITIEVQATNVFVGQDRESLAI	VN	640	
Db	512	DGWTMTIKVSYSS--GSANVOVDV	S	534	

RESULT 14  
 US-08-894-818B-5  
 Sequence 5, Application US/08894818B  
 Patent No. 6261822  
 GENERAL INFORMATION:  
 APPLICANT: TAKAKURA, HiKaru  
 APPLICANT: MORISHITA, Mio  
 APPLICANT: YAMAMOTO, Katsuhiko  
 APPLICANT: MITTA, Masamori  
 APPLICANT: ASADA, Kiyozo  
 APPLICANT: TSUNASAMA, Susumu  
 APPLICANT: KATO, Ikunoshin  
 TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES  
 NUMBER OF SEQUENCES: 42  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Brody and Nelmark  
 STREET: 419 Seventh Street N.W., Ste. 300  
 City: Washington  
 STATE: D.C.  
 COUNTRY: United States of America  
 Zip: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/894,818B  
 FILING DATE: 20-MAY-1998  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/JP96/03253  
 FILING DATE: 07-NOV-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 323285/1995  
 FILING DATE: 12-DEC-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brody, Roger L.  
 REGISTRATION NUMBER: 25,618





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 1, 2003, 12:41:04 ; Search time 24.0188 Seconds  
(without alignments)  
1629.021 Million cell updates/sec

Title: US-09-920-954-2

Perfect score: 3102  
Sequence: 1 MRKRKVFSLVLSAAALST.....EVOQVNPVGPQKFLSLAIVN 640

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications\_Aa:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEM\_PUB.pdp:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCOT\_NEM\_PUB.pdp:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEM\_PUB.pdp:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEM\_PUB.pdp:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pdp:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pdp:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCOTUS\_PUBCOMB.pdp:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pdp:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEM\_PUB.pdp:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pdp:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEM\_PUB.pdp:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pdp:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEM\_PUB.pdp:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pdp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2155	69.5	434	9	US-09-985-689A-1
2	2155	69.5	434	9	US-09-985-689A-2
3	2082	67.1	434	9	US-09-985-689A-6
4	2060.5	66.4	433	9	US-09-985-689A-7
5	1952.5	62.9	433	9	US-09-985-689A-5
6	1948.5	62.8	433	9	US-09-985-689A-3
7	1941.5	62.6	433	9	US-09-985-689A-4
8	451.5	14.6	659	12	US-10-090-624-12
9	408	13.2	654	12	US-10-090-624-16
10	391	12.6	652	12	US-10-090-624-1
11	391	12.6	522	12	US-10-090-624-4
12	307.5	9.9	418	10	US-09-966-921A-2
13	299.5	9.7	1398	12	US-10-090-624-6
14	280.5	9.0	595	9	US-09-927-827-55
15	272	8.8	397	9	US-09-779-334A-5
16	271.5	8.8	379	9	US-09-813-408-6
17	270	8.7	380	9	US-09-927-827-55
18	250.5	8.1	380	9	US-09-824-893A-261
19	250	8.1	377	9	US-09-813-408-1

20	250	8.1	382	9	US-09-813-408-7	Sequence 7, Appl1
21	246	7.9	271	9	US-09-813-408-2	Sequence 2, Appl1
22	245	7.9	382	12	US-10-090-624-11	Sequence 31, Appl1
23	243	7.8	379	9	US-09-813-408-13	Sequence 13, Appl1
24	242.5	7.8	379	9	US-09-813-408-10	Sequence 10, Appl1
25	242	7.8	380	9	US-09-813-408-19	Sequence 19, Appl1
26	240.5	7.8	380	9	US-09-813-408-16	Sequence 16, Appl1
27	239	7.7	269	9	US-09-873-139-1	Sequence 14, Appl1
28	239	7.7	269	9	US-09-976-414-8	Sequence 8, Appl1
29	239	7.7	269	10	US-09-837-235-16	Sequence 16, Appl1
30	239	7.7	269	10	US-09-060-854B-6	Sequence 6, Appl1
31	239	7.7	269	12	US-10-075-907-1	Sequence 1, Appl1
32	239	7.7	269	12	US-10-075-895-1	Sequence 1, Appl1
33	238.5	7.7	381	10	US-09-920-118-16	Sequence 16, Appl1
34	237.5	7.7	379	9	US-09-813-408-11	Sequence 11, Appl1
35	235.5	7.6	379	10	US-09-920-118-14	Sequence 14, Appl1
36	235	7.6	269	9	US-09-813-408-4	Sequence 4, Appl1
37	234.5	7.6	1150	10	US-09-870-122-3	Sequence 3, Appl1
38	234.5	7.6	1181	10	US-09-870-122-15	Sequence 15, Appl1
39	234	7.5	310	9	US-09-813-408-15	Sequence 23, Appl1
40	232	7.5	269	9	US-09-813-408-5	Sequence 5, Appl1
41	231.5	7.5	1167	10	US-09-870-122-2	Sequence 2, Appl1
42	230.5	7.4	279	9	US-09-813-408-21	Sequence 21, Appl1
43	229	7.4	269	9	US-09-779-334A-7	Sequence 7, Appl1
44	228	7.4	266	10	US-09-837-235-18	Sequence 18, Appl1
45	228	7.4	378	9	US-09-813-408-14	Sequence 14, Appl1

## ALIGNMENTS

RESULT 1  
US-09-985-689A-1  
; Sequence 1, Application US/09985689A  
; Publication No. US20030022351A1  
; GENERAL INFORMATION:  
; APPLICANT: HATADA, YUJI  
; APPLICANT: OGAWA, AKINORI  
; APPLICANT: KAGEYAMA, YASUSHI  
; APPLICANT: SATO, TSYOSHI  
; APPLICANT: ARAKI, HIROYUKI  
; APPLICANT: SUMITOMO, NOBUYUKI  
; APPLICANT: OKUDA, MITSUYOSHI  
; APPLICANT: SAEKI, KATSUHIKA  
; TITLE OF INVENTION: Alkaline proteases  
; FILE REFERENCE: 215483US0  
; CURRENT APPLICATION NUMBER: US/09/985,689A  
; CURRENT FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: JP P2000-355166  
; PRIOR FILING DATE: 2000-11-22  
; PRIOR APPLICATION NUMBER: JP P2001-114048  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 434  
; TYPE: PRT  
; ORGANISM: Bacillus sp.  
US-09-985-689A-1

Query Match 69.5% Score 2155 DB 9 Length 434

Best local Similarity 96.3% Pred No 5 3e-175

Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 207 NDVAKGIVADYADYAGSYGLYKGGGIVAVADTGLDTRGRSSMHEAFKRTITALLAALGRN 266  
DB 1 NDVAKGIVADYADYAGSYGLYKGGGIVAVADTGLDTRGRSSMHEAFKRTITALLAALGRN 60  
QY 267 NADNTGHTGTHVAGSVLGLNGXTNKGMAPOANLFGOSIMDSXGGLGLPSNLQTLFSSQAYS 326  
DB 61 NADNTGHTGTHVAGSVLGLNGXTNKGMAPOANLFGOSIMDSXGGLGLPSNLQTLFSSQAYS 120  
QY 327 AGARHTNSGCAVAGATYTDTSRNVDDYVRKNDMTLFLFAGNEXPNCGTISAPDTAKNAI 386

```
Db 121 AGARLHTNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGNGGTISAPGTAKNAI 180
Qy 387 TVGATENLRPSFGSYADININHAOFSSRGPTKDGRIKPDVMAPGTXIISARSLAPDSSF 446
Db 181 TVGATENLRPSFGSYADININHAOFSSRGPTKDGRIKPDVMAPGFTIISARSLAPDSSF 240
Qy 447 WANHDSKYAYMGSTMATPIVAGNVAQLREHFVKNRGITTPKPSLKALLAGADXLGY 506
Db 241 WANHDSKYAYMGSTMATPIVAGNVAQLREHFVKNRGITTPKPSLKALLAGADXLGY 300
Qy 507 PNGNGMGRVTLDKSLNAYVNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 566
Db 301 PNGNGMGRVTLDKSLNAYVNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 360
Qy 567 SVTLVNDLDVITAPNGTYXVGNDFXPPXXNMWDRNNEVFIXXPOSGTYTIEVOAYN 626
Db 361 SVTLVNDLDVITAPNGTYXVGNDFXPPXXNMWDRNNEVFIXXPOSGTYTIEVOAYN 420
Qy 627 VPGVQXFSLAIVN 640
Db 421 VPGVQXFSLAIVN 434
```

```
RESULT 2
US-09-985-689A-2
; Sequence 2, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-2
```

Query Match 69.5%; Score 2155; DB 9; Length 434;  
Best Local Similarity 96.3%; Pred. No. 5.3e-175;

Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy 207 NDVARGIKADVAAOSSYGLYGOGIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 266
Db 1 NDVARGIKADVAAOSSYGLYGOGIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Qy 267 NANDNGHGTIVAGSVLNGXTNKGMAPQANLVFOSIMDSXGGLGRLPSNIQTLEFSQAXS 326
Db 61 NANDNGHGTIVAGSVLNGXTNKGMAPQANLVFOSIMDSXGGLGRLPSNIQTLEFSQAXS 120
Qy 327 AGARLHTNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGNGGTISAPGTAKNAI 386
Db 121 AGARLHTNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGNGGTISAPGTAKNAI 180
Qy 387 TVGATENLRPSFGSYADININHAOFSSRGPTKDGRIKPDVMAPGTXIISARSLAPDSSF 446
Db 181 TVGATENLRPSFGSYADININHAOFSSRGPTKDGRIKPDVMAPGTXIISARSLAPDSSF 240
```

```
Qy 447 WANHDSKYAYMGSTMATPIVAGNVAQLREHFVKNRGITTPKPSLKALLAGADXLGY 506
Db 241 WANHDSKYAYMGSTMATPIVAGNVAQLREHFVKNRGITTPKPSLKALLAGADXLGY 300
Qy 507 PNGNGMGRVTLDKSLNAYVNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 566
Db 301 PNGNGMGRVTLDKSLNAYVNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 360
Qy 567 SVTLVNDLDVITAPNGTYXVGNDFXPPXXNMWDRNNEVFIXXPOSGTYTIEVOAYN 626
Db 361 SVTLVNDLDVITAPNGTYXVGNDFXPPXXNMWDRNNEVFIXXPOSGTYTIEVOAYN 420
Qy 627 VPGVQXFSLAIVN 640
Db 421 VPGVQXFSLAIVN 434
```

```
RESULT 3
US-09-985-689A-6
; Sequence 6, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-6
```

Query Match 67.1%; Score 2082; DB 9; Length 434;  
Best Local Similarity 91.5%; Pred. No. 8.3e-169;

Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

```
Qy 207 NDVARGIKADVAAOSSYGLYGOGIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 266
Db 1 NDVARGIKADVAAOSSYGLYGOGIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Qy 267 NANDNGHGTIVAGSVLNGXTNKGMAPQANLVFOSIMDSXGGLGRLPSNIQTLEFSQAXS 326
Db 61 NANDNGHGTIVAGSVLNGXTNKGMAPQANLVFOSIMDSXGGLGRLPSNIQTLEFSQAXS 120
Qy 327 AGARLHTNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGNGGTISAPGTAKNAI 386
Db 121 AGARLHTNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGNGGTISAPGTAKNAI 180
Qy 387 TVGATENLRPSFGSYADININHAOFSSRGPTKDGRIKPDVMAPGTXIISARSLAPDSSF 446
Db 181 TVGATENLRPSFGSYADININHAOFSSRGPTKDGRIKPDVMAPGFTIISARSLAPDSSF 240
Qy 447 WANHDSKYAYMGSTMATPIVAGNVAQLREHFVKNRGITTPKPSLKALLAGADXLGY 506
Db 241 WANHDSKYAYMGSTMATPIVAGNVAQLREHFVKNRGITTPKPSLKALLAGADXLGY 300
Qy 507 PNGNGMGRVTLDKSLNAYVNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 566
Db 301 PNGNGMGRVTLDKSLNAYVNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA 360
```



```

; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUOSHI
; APPLICANT: SAKKI, KATSUHISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 433
; TYPE: prt
; ORGANISM: Bacillus sp.
; US-09-985-689A-3

Query Match          62.8%; Score 1948.5; DB 9; Length 433;
Best Local Similarity 86.4%; Pred. No. 1.8e-157;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

QY 207 NDVARGIKADYAOSSYGLYGQGIYAAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 266
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 1 NDVARGIKADYAOSSYGLYGQGIYAAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60

QY 267 NANDRNGHGTTHAGSVLNGXTNKGMAPOANLVFOSIMDSXGLGLPSNLTLEFSQAXS 326
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 61 NANDRNGHGTTHAGSVLNGXTNKGMAPOANLVFOSIMDSXGLGLPSNLTLEFSQAXS 119

QY 327 AGARIHTNSWGAHVNGAYTTDSRNDVYRKNDMTLFAAGNEXPNGGTISAPGTAKNAI 386
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 120 AGARIHTNSWGAHVNGAYTTDSRNDVYRKNDMTLFAAGNEXPNGGTISAPGTAKNAI 179

QY 387 TYGATEENLRPSFGSYADINNHVAFSSRGPTDGRKRPDYMAPGTXIISARSSLAPDSSF 446
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 180 TYGATEENLRPSFGSYADINNHVAFSSRGPTDGRKRPDYMAPGTXIISARSSLAPDSSF 239

QY 447 WANHSKYAYMGCTSMATPIYAGNVAOLREHFVNKRGITPKPSLKAALIGAADXGLGY 506
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 240 WANHSKYAYMGCTSMATPIYAGNVAOLREHFVNKRGITPKPSLKAALIGAADXGLGY 299

QY 507 PNGNOGWRVTLDKSLNVAAYVNESSXLSSTQKATYXFTATAGKPLKISLWSDAPASTTA 566
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 300 PNGNOGWRVTLDKSLNVAAYVNESSXLSSTQKATYXFTATAGKPLKISLWSDAPASTTA 359

QY 567 SYTLVNDLDTLTAPNGTYVGNDFXPPXXMWDGRNNVENFVIXXPGSGTITIEVOAYN 626
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 360 SYTLVNDLDTLTAPNGTYVGNDFXPPXXMWDGRNNVENFVIXXPGSGTITIEVOAYN 419

QY 627 VPGPQXESLAIYN 640
    |||||.....:
DB 420 VPGPQXESLAIYN 433

RESULT 7
US-09-985-689A-4
; Sequence 4, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
```

```

; APPLICANT: OKUDA, MITSUOSHI
; APPLICANT: SAKKI, KATSUHISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 433
; TYPE: prt
; ORGANISM: Bacillus sp.
; US-09-985-689A-4

Query Match          62.6%; Score 1941.5; DB 9; Length 433;
Best Local Similarity 86.2%; Pred. No. 7e-157;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 207 NDVARGIKADYAOSSYGLYGQGIYAAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 266
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 1 NDVARGIKADYAOSSYGLYGQGIYAAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60

QY 267 NANDRNGHGTTHAGSVLNGXTNKGMAPOANLVFOSIMDSXGLGLPSNLTLEFSQAXS 326
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 61 NANDRNGHGTTHAGSVLNGXTNKGMAPOANLVFOSIMDSXGLGLPSNLTLEFSQAXS 119

QY 327 AGARIHTNSWGAHVNGAYTTDSRNDVYRKNDMTLFAAGNEXPNGGTISAPGTAKNAI 386
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 120 AGARIHTNSWGAHVNGAYTTDSRNDVYRKNDMTLFAAGNEXPNGGTISAPGTAKNAI 179

QY 387 TYGATEENLRPSFGSYADINNHVAFSSRGPTDGRKRPDYMAPGTXIISARSSLAPDSSF 446
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 180 TYGATEENLRPSFGSYADINNHVAFSSRGPTDGRKRPDYMAPGTXIISARSSLAPDSSF 239

QY 447 WANHSKYAYMGCTSMATPIYAGNVAOLREHFVNKRGITPKPSLKAALIGAADXGLGY 506
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 240 WANHSKYAYMGCTSMATPIYAGNVAOLREHFVNKRGITPKPSLKAALIGAADXGLGY 299

QY 507 PNGNOGWRVTLDKSLNVAAYVNESSXLSSTQKATYXFTATAGKPLKISLWSDAPASTTA 566
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 300 PNGNOGWRVTLDKSLNVAAYVNESSXLSSTQKATYXFTATAGKPLKISLWSDAPASTTA 359

QY 567 SYTLVNDLDTLTAPNGTYVGNDFXPPXXMWDGRNNVENFVIXXPGSGTITIEVOAYN 626
    |||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
DB 360 SYTLVNDLDTLTAPNGTYVGNDFXPPXXMWDGRNNVENFVIXXPGSGTITIEVOAYN 419

QY 627 VPGPQXESLAIYN 640
    |||||.....:
DB 420 VPGPQXESLAIYN 433

RESULT 8
US-10-090-624-12
; Sequence 12, Application US/10090624
; Patent No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshih
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090, 624
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445, 472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
```

;; NUMBER OF SEQ ID NOS: 33  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 12  
;; LENGTH: 659  
;; TYPE: PRT  
;; ORGANISM: Thermococcus celer  
US-10-090-624-12

Query Match 14.6%; Score 451.5; DB 12; Length 659;  
Best Local Similarity 25.9%; Pred. No. 3,6e-30;  
Matches 157; Conservative 92; Mismatches 207; Indels 151; Gaps 24;

QY 70 LKRLKLETPVANKKLIHFNPILEETK--QXLEXTGAILIDYIDYIYEVGSD 127  
DB 43 LTPLEFKVVRMMNOEDVTVMFSTYGRRAVKVLRMGAYQ-----KYSIK----- 92  
QY 128 VXSXXXIIEHVESVEPYLPXYIIDPOLFTKGASXLVKAALDTRKXNKVEQLRIGETIAQ 187  
DB 93 -----IYAVAVK-----IKARDLLLAGMIDIDYFG--WTRVSGIKFTQE 130  
QY 188 XXXSNDVXYITAKPEYKVMNDVARGI--VKADVAOSSYGLYGQQLVAVADLTGTGRND 245  
DB 131 -----DYKQVDDATSVSQIGADYVWNSLIGDGSQVVAIVDTGIDAN--- 173  
QY 246 SSMHEARCKITALY-ALGRTNNANDTNGHGTIVAGSVLGNXGXTNK--GMAPOANLVFQ 301  
DB 174 ---HPDLKGVIGWDVAVNGRSTPYDDGCHGTIVAGTGSVNGQYIGVAPGAKLVG 230  
QY 302 SIM--DSXGIGLGLPSLQTLFSQAXSAGARI-----HTNSMGAVNCAXTTD 347  
DB 231 KVLGADSGSVSTINGVDMVYQNKDKGIVINLSGSSOSSDGDLSQAVNNAKMA- 289  
QY 348 SRNVDVYRKNDMTILFAAGNEXPNCGTISAPGTAKNAITVGTETLRPSFGSYDNTNH 407  
DB 230 -----GLVYVVAAGNSGPNFTVYGSPPAASKVITVGA-----VDSMDN 327  
QY 408 VAOFSSGPTKDGRIKPDVMAFGTXILSARSSLAPDSSFFMANHDSKYAYMGSTMAFPY 467  
DB 328 IASSSSGPPADRLKPEVAVAGVDIIAPRAS--GTSMTGPIINDYTTASGTSMAFHV 384  
QY 468 AGNVAQLREHFVNKRITP--KPSLLKAA-LIAGAAXGLGYPNQNGMGVTLDKSL- 522  
DB 385 SGVALTLQ---AHPMTDPDKYKTALLETADIVAPKEIADIAV-----GAGHVNYKKA 436  
QY 523 --NVAAYNESSXLSSTOKATYFTATACKPLKISLVMSDAPASTASVTLVNDLDTVTA 580  
DB 437 YDYAKLITFTGSVADKSAFHTFDVSGATFYTATLYND-----TGSSDIDLTYLD 486  
QY 581 PNG-----TXYVGNDFXPPXXNMWDRNNVEVFINXPQSTYTIIEVQATNVVGPQX 633  
DB 487 PNGNEVDYSTIAYG-----FEKVGTYNFTAGTWIVKVVSTK---GAAN 527  
QY 634 FSLAIYN 640  
DB 528 YQDVVS 534

RESULT 9  
US-10-090-624-16  
; Sequence 16, Application US/10090624  
; Patent No. US20020132335A1  
; GENERAL INFORMATION:  
; APPLICANT: TAKAKURA, Hikaru  
; APPLICANT: MORISHITA, Mio  
; APPLICANT: SHIMOJO, Tomoko  
; APPLICANT: ASADA, Kiyozo  
; APPLICANT: KATO, Ikunoshi  
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE  
; FILE REFERENCE: TAKAKURA-6  
; CURRENT APPLICATION NUMBER: US/10/090, 624  
; PRIOR APPLICATION NUMBER: 09/445, 472  
; PRIOR FILING DATE: 1999-12-06

;; PRIOR APPLICATION NUMBER: 151969/1997  
;; PRIOR FILING DATE: 1997-06-10  
;; NUMBER OF SEQ ID NOS: 33  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 16  
;; LENGTH: 654  
;; TYPE: PRT  
;; ORGANISM: Pyrococcus furiosus  
US-10-090-624-16

Query Match 13.2%; Score 408; DB 12; Length 654;  
Best Local Similarity 25.9%; Pred. No. 1,8e-26;  
Matches 162; Conservative 80; Mismatches 210; Indels 174; Gaps 26;

QY 63 VESENVKLLKKKLETPVANKKLIH--IKQFNPILEETKQXLEXTGAKLIIDYIDYAY 120  
DB 35 VEKNYGLTPTGLERKIQKLNPEEISTYIYFENHREKELAVRLELMGAKY-----RY 87  
QY 121 IVEYEGDVXSXXXIIEHVESVEPYLPXYIIDPOLFTKGASXLVKAALDTRKXNKVEQLR 180  
DB 88 V-----YHIIPAI-----AADLKYRDLVYISGLTGKAKLS 118  
QY 181 GEXIAQXXXSNDVXYITAKPEYKVMNDVARGIYKADVAOSSYGLYGQQLVAVADLTGLD 240  
DB 119 GVRPIQEDYK-----VYSALEGLDESAQVAVATYVWMLGYD--GSGITIGIITGID 170  
QY 241 TGRNDSMHEARCKITALYALGRTNNAN-----DTNGHGTIVAGSVLGNXGXTNK---K 290  
DB 171 ---ASHPDLQGV-----IGWDFPNGRSRYDDHGHGTIVASIAAGTGAASNGKRY 219  
QY 291 GMAPOANLVFQSIM--DSXGIGLGLPSLQTLFSQAXSAGARIHTNSMGA----- 338  
DB 220 GMAAGALACIKVLAGDGSISTITIKGEVANDKDKGIGKIVINLSGSSOSSDGDLDAL 279  
QY 339 --AVNCAVTTDSRNVDVYRKNDMTILFAAGNEXPNCGTISAPGTAKNAITVGTENLRP 396  
DB 280 SQAVNAAMDA-----GLVYVVAAGNSGPNKTYTIGSPAAASKVITVGA----- 321  
QY 397 SFGSYDNTNHVAFSSRQPTKDGRIKPDVMAFGTXILSARSSLAPDSSFFMANHDSKYAY 456  
DB 322 ---VDKYVITISFSSRPADRLKPEVAVAPQWIIAPRAS--GTSMGOPINDYTTA 373  
QY 457 MGSMTAPTPVAGNVAQLREHFVNKRITP--PSLLKALLIGA-----ADXGIGY 506  
DB 374 APGSMATPPIVAGIALALLQ-----AHPMTDPDKYKTALLETADIVKPEIAD--IAY 424  
QY 507 PNGMGWGRVTLDKSLNVAAYNESSXLSSTOKA---TYXFTATAGKRLKISLVMSDAP 561  
DB 425 ---GAGRNATKAIN--YDNVAKLVFTGVANKSGTHQFVISCASVYRTILYWDNAN 477  
QY 562 ASTTASVTLVNDLDTYITAPNG-----TXYVGNDFXPPXXNMWDRNNVEVFINXPQ 614  
DB 478 -----SDLDLTLDPNGNOVDYSTIAYTG-----FEKVGTYNFT 511  
QY 615 SGTYTIEVQATNVVGPQXFSLAIYN 640  
DB 512 DGTWITIKVVSYS---GSANYQDVVS 534

RESULT 10  
US-10-090-624-1  
; Sequence 1, Application US/10090624  
; Patent No. US20020132335A1  
; GENERAL INFORMATION:  
; APPLICANT: TAKAKURA, Hikaru  
; APPLICANT: MORISHITA, Mio  
; APPLICANT: SHIMOJO, Tomoko  
; APPLICANT: ASADA, Kiyozo  
; APPLICANT: KATO, Ikunoshi  
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE  
; FILE REFERENCE: TAKAKURA-6  
; CURRENT APPLICATION NUMBER: US/10/090, 624  
; PRIOR FILING DATE: 2002-03-06

; PRIOR APPLICATION NUMBER: 09/445,472  
; PRIOR FILING DATE: 1999-12-06  
; PRIOR APPLICATION NUMBER: 151969/1997  
; PRIOR FILING DATE: 1997-06-10  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 412  
; TYPE: PRT  
; ORGANISM: Pyrococcus furiosus  
US-10-090-624-1

Query Match 12.6%; Score 391; DB 12; Length 412;  
Best Local Similarity 28.9%; Pred. No. 2.6e-25;  
Matches 134; Conservative 56; Mismatches 145; Indels 128; Gaps 20;

QY 224 GYGGGQIYAVADTGLDGRNDSMHEAFRGKITALYALGRNNNN-----DTNGHGH 277  
DB 22 GYDGGSGITIGLIDTGLD-----ASHPDLQGRV-----IGWDFVNGRSYPYDDHGHGH 70  
QY 278 VAGSVLGNQXTN-----KGMAPQANLVFOSIM--DSXGGLGGLPSNLQTLFSAQXSAGARI 331  
DB 71 VASIAAGTGAASNGRYKKGAPGAKLAGIKVLGADSGSISTIIKGVEMAVDMKKDKYGIKV 130  
QY 332 HTNSMGA-----AVNGAYTDSRNVDYVRKNDMTILPAGNEXNGGTTISAP 379  
DB 131 INLSLSSQSSDGTDLALSOAVNAAMD-----GLVVVVAAGNSGPNKRYTIGSP 178  
QY 380 GTAKNAITVYGAENLRPSGSGYADNINHYAOFSSRGPRFKDGRKPDVMAFGTXIISARSS 439  
DB 179 AAASKVITYGA-----VDKYDYITSSSSKGPDLADGLKEVVAAPGWTIAARAS 227  
QY 440 LAPDSSFMANHDSKYAYMGTSMATPIVAGNVAOLREHFVKNRGITPK--PSILKAALIA 497  
DB 228 ---GTSMGQPINDYTTAAGTSMATPHVAGIAALLQ-----AHPSTPKVKTALIE 277  
QY 498 GA-----ADXGLGYPNGNGKGRVITLDSKLNAYVNESSXLSLSOKA-----TYXFT 544  
DB 278 TADIYKPEIAD--IAY-----GAGRVAAYKAIN--YDVAKLVTGYVANKGSOIHOFV 328  
QY 545 ATAGKPLKISLYSDAPASTASVTLVNDLDEVITAPNG-----TXYVGNDFFXPXX 597  
DB 329 ISGASFVATITLYMDAN-----SDLDLYLDPNGNOVDYSTIATYG----- 369  
QY 598 NMDGRNNEVNFINKPQSGTYTIEVOAYNVVPQXFSIAVIN 640  
DB 370 -----FEKVGYNPTDGTWTIKVVSYS--GSANYQVDVVS 402

RESULT 11  
US-10-090-624-4  
; Sequence 4, Application US/10090624  
; Patent No. US2002012335A1  
; GENERAL INFORMATION:  
; APPLICANT: TAKAKURA, Hikaru  
; APPLICANT: MORISHITA, Mio  
; APPLICANT: SHIMOJO, Tomoko  
; APPLICANT: ASADA, Kiyozo  
; APPLICANT: KATO, Ikunoshin  
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE  
; FILE REFERENCE: TAKAKURA-6  
; CURRENT APPLICATION NUMBER: US/10/090,624  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: 09/445,472  
; PRIOR FILING DATE: 1999-12-06  
; PRIOR APPLICATION NUMBER: 151969/1997  
; PRIOR FILING DATE: 1997-06-10  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 522  
; TYPE: PRT  
; ORGANISM: Pyrococcus furiosus

; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: (428)..(428)  
; OTHER INFORMATION: Xaa at position 428 is Gly or Val.  
US-10-090-624-4

Query Match 12.6%; Score 391; DB 12; Length 522;  
Best Local Similarity 28.9%; Pred. No. 3.6e-25;  
Matches 134; Conservative 56; Mismatches 145; Indels 128; Gaps 20;

QY 224 GYGGGQIYAVADTGLDGRNDSMHEAFRGKITALYALGRNNNN-----DTNGHGH 277  
DB 22 GYDGGSGITIGLIDTGLD-----ASHPDLQGRV-----IGWDFVNGRSYPYDDHGHGH 70  
QY 278 VAGSVLGNQXTN-----KGMAPQANLVFOSIM--DSXGGLGGLPSNLQTLFSAQXSAGARI 331  
DB 71 VASIAAGTGAASNGRYKKGAPGAKLAGIKVLGADSGSISTIIKGVEMAVDMKKDKYGIKV 130  
QY 332 HTNSMGA-----AVNGAYTDSRNVDYVRKNDMTILPAGNEXNGGTTISAP 379  
DB 131 INLSLSSQSSDGTDLALSOAVNAAMD-----GLVVVVAAGNSGPNKRYTIGSP 178  
QY 380 GTAKNAITVYGAENLRPSGSGYADNINHYAOFSSRGPRFKDGRKPDVMAFGTXIISARSS 439  
DB 179 AAASKVITYGA-----VDKYDYITSSSSKGPDLADGLKEVVAAPGWTIAARAS 227  
QY 440 LAPDSSFMANHDSKYAYMGTSMATPIVAGNVAOLREHFVKNRGITPK--PSILKAALIA 497  
DB 228 ---GTSMGQPINDYTTAAGTSMATPHVAGIAALLQ-----AHPSTPKVKTALIE 277  
QY 498 GA-----ADXGLGYPNGNGKGRVITLDSKLNAYVNESSXLSLSOKA-----TYXFT 544  
DB 278 TADIYKPEIAD--IAY-----GAGRVAAYKAIN--YDVAKLVTGYVANKGSOIHOFV 328  
QY 545 ATAGKPLKISLYSDAPASTASVTLVNDLDEVITAPNG-----TXYVGNDFFXPXX 597  
DB 329 ISGASFVATITLYMDAN-----SDLDLYLDPNGNOVDYSTIATYG----- 369  
QY 598 NMDGRNNEVNFINKPQSGTYTIEVOAYNVVPQXFSIAVIN 640  
DB 370 -----FEKVGYNPTDGTWTIKVVSYS--GSANYQVDVVS 402

RESULT 12  
US-09-966-921A-2  
; Sequence 2, Application US/09966921A  
; Patent No. US20020103100A1  
; GENERAL INFORMATION:  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Lassen, Soren  
; APPLICANT: Pedersen, Poulsen  
; TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides Having Proteolytic Activity  
; FILE REFERENCE: 10097,200-US  
; CURRENT APPLICATION NUMBER: US/09/966,921A  
; CURRENT FILING DATE: 2001-09-28  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 418  
; TYPE: PRT  
; ORGANISM: Bacillus sp.  
US-09-966-921A-2

Query Match 9.9%; Score 307.5; DB 10; Length 418;  
Best Local Similarity 28.4%; Pred. No. 3.3e-18;  
Matches 116; Conservative 57; Mismatches 154; Indels 81; Gaps 20;

QY 140 SVEPLPYXIDPQLFTGASXLVAKXALDTQXKKEVQLRIEXIAQXXSNDVXYITA 199  
DB 49 SISTFOSQYDVAWMDGKGEFTTIDVDAKQQLQTLQSKNDIQIKVNMETVETTEKAETYA 108  
QY 200 KPEYK-----VMDVARGIVKADVAQSSYGLYGOGQIVAAVADTGLDGRNDSMHEAF 252



```
Db 109 VPSOTFWGKISLYND--OSITKT-----TGGSGIKVAVLDTGCVYTS-----HLDL 152
Oy 253 RGTITALYALGRTN-----NANDTNGHGTAVAGSVLNGXTN-----KGMAPQANL-VFQS 302
Db 153 AGSAEOCKDFPTQSNPLVDGSCFDRQGHGTIVLAHAGSGNSOGYVAVPAQAKLMAYK 212
Oy 303 IMDSXGIG-GLPSNLQTLFESQXSAGARIHTN-SWGAANGVATYTDNRNVDYRKNDM 360
Db 213 LGDNGSYSDIDIAAIRHVADEASRTGSKYVINNSLSSAKDILLN---SAVDYAGKGV 269
Oy 361 TILFAAGNEKPNNGTISAGCTAKNAITGATENLRPSPGSAONINHVQFSSRG-PTKD 419
Db 270 LIVAAAGNSGSGNTTIGFPGGLYNVAVALENVOQN-GT-----RVADFSRGRPATFA 323
Oy 420 G-----RIKPDVAPGTXLISARSSLADSSFWANHDSKTAIYNGTSMAPIYAGNVA-- 472
Db 324 GDYIIOERDIEVSAPGASV-----ESTW--YTGGYNTISGTSMAPIPVAGLAKI 371
Oy 473 -----QREHFVNKRGITPPRSLKKAALIAAAD---XGLGYP 507
Db 372 WSANTSLSHSQLRTE-LQNRRA---KYVDIKGIGAGTGDDYASGFGYP 415

RESULT 13
US-10-090-624-6
; Sequence 6, Application US/10090624
; Patent No. US2002013335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, HiKaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 1398
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-6

Query Match 9.7%; Score 299.5; DB 12; Length 1398;
Best Local Similarity 22.9%; Pred. No. 7.9e-17;
Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;

Oy 97 LEETKOLEXTGAKILD-----YIPDAIYVEE-----GDYXSKXXXIE 136
Db 77 LEETKTELEKGAELIDENRVLNMLLVKIKPEKVELNYISLEKAMLRKVKLSPPIVE 136
Oy 137 -HVESVEPYLPXXYIDPOLFT-----KGASXLVAKAALDTKQXNEV 177
Db 137 KDVTKRPE-----SLEPKMYNSTWVINALQFIOEGYDGGVVVAALDGTVPNHPFLSI 191
Oy 178 QLRGIEIXIAQXXXNDVXYITAKPEY-KVMN-----DVARGI----- 213
Db 192 TPDGRRKIIEMKPFDEGFDVTSFSFSKYVNGTLINTTFOVASGLTLNESGCLMEYVK 251
Oy 214 -----VKADVAOSSYGLY-----GQOQIYAAVADT 237
Db 252 TVYVSANTIGNITSANGIYHFGILPERYFDLFDGDEDEYVVLVNSNGNGYDIAAYDT 311
Oy 238 GLDTGRNDS-----SMHEAFRGKITALXALGRNNAN-----DTNGHGTIYAG 280
Db 312 DLDVDFIDEPFLQGVNTYTDVAVPFSYITGPLNYLAEIDRGEYAVFGMDGHHGHIVAG 371
```

```
Oy 281 SVLNGXTN-----KGMAPQANLFGOSIMDSXG 308
Db 372 TVAGYDSNDNDAMDMLSMYSGEWEFSLRYGMDYTNVTTPTDVGVAPOAIMAIVHLS-D 430
Oy 309 GLGLPSNLQTLFESQXSAGARIHTNSGAAVNGAYT--TDSRN--VDDYRKNDMTILF 364
Db 431 GRGSMMDIIEGM-TYAAHTPGADVISMISLGG--NAPLYDGTDPDESVAVDLEIKYGVFVI 487
Oy 365 AAGREKPNNGTISAGCTAKNAITGATENLRPSPGSAONINHVQFSSRG-PTKD 419
Db 488 AAGNEKPNNGTISAGCTAKNAITGATENLRPSPGSAONINHVQFSSRG-PTKD 419
Oy 407 HVAOFSSRGPTKDGRIKPDVAPGTXLISARSSLADSSFWANHDSKTAIYNGTSMAPI 466
Db 546 RIAPFSSRGPRIDGERTKPVVAPGITYSLPMIIGADP-----MSGTSMAPIH 595
Oy 467 VAGNVAOLREHFVNKRGITPPRSLKKAALIAAAD---DXGLGYPNGOGMGVYTLDK 520
Db 596 VSGVVALISG-AKAGELYNPDIIKVLSEGATWLECDPYTGOKYTELDGHLGVNVTK 654
Oy 521 SLNVAAYNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTTA---SVTLVNDLVL 577
Db 655 SMEI-----LKAINGTTLPIVDHMDKSYSDFAEYLGVDVIRGLYAR 696
Oy 578 ITAPN-----GTXYVGN---DFXXXKXXNM-----DG-----RNRYE----- 606
Db 697 NSIPDIEMHIKYVDGTERTEIYATEPMIKPFVSGSVILENTEFVLARKYVDEGLEP 756
Oy 607 -----WFTINXQSGTYTTEVOATNVPYGPQXS 635
Db 757 GLIYGRILIIDDP--TPVLEDELMTIYIPEKFI 788

RESULT 14
US-09-927-827-59
; Sequence 59, Application US/09927827
; Publication No. US20030036176A1
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Ramseller, Thomas M.
; TITLE OF INVENTION: Directed Genetic Engineering of Xanthomonas campestris
; FILE REFERENCE: 38-10(15824)B
; CURRENT APPLICATION NUMBER: US/09/927,827
; CURRENT FILING DATE: 2001-08-10
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 69
; SEQ ID NO 59
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-09-927-827-59

Query Match 9.0%; Score 280.5; DB 9; Length 595;
Best Local Similarity 24.9%; Pred. No. 1e-15;
Matches 125; Conservative 64; Mismatches 163; Indels 151; Gaps 21;

Oy 93 NGPILBETKOLEXTGAKILDYIPDAIYVEEYEGDYXSKXXXIEHVESVEPYLPXXYIDP 152
Db 136 NAVILIKLSLEQLEAAGA-----DYRHEVLIELD-----RRVLGADP---AGVRA 176
Oy 153 QLFKRGASXLVAKAALDTKQXNEKVEOLRG-----IEIXIAQXXXNDVXYIT 198
Db 177 QLTIHMQIAGHSTGDLGDALDEAIELEDSLQRLISARLRLREIETLQSRRLAGPFRRW 236
Oy 199 AKPEYKVMNDVAGIYKADVAOSSYGLYCGOQIYAVADPGLDTGRNDSMEAF-----R 253
Db 237 RDAGKRLIHVSQNVLIHDAARTAYBARGOOIGHAVLDGT-----AAAHPRFVYGER 290
Oy 254 GKITALY-----ALGRTNNAANDTNGHGTIYAGSVLG-----NGXTNK- 290
Db 291 DNVAQMDCTRRGAPKRLTRADGKRFETALDGHGHTIYAGIITAGCCRAAVPDASGKPGEL 350
```



GenCore version 5.1.4-p5\_A578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:36:24 ; Search time 17.5137 Seconds  
(without alignments)  
3513.026 Million cell updates/sec

Title: US-09-920-954-2

Sequence: 1 MRKKKVFSLVSLAAALST.....EQVQNVNPGVQKSLAIYN 640

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1.\*  
2: p1r2.\*  
3: p1r3.\*  
4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518.5	16.7	1743	2 T18279	multidrug resistan
2	477.5	15.4	1905	2 T18267	multidrug resistan
3	350.5	11.3	444	2 B83891	intracellular alka
4	319.5	10.3	442	2 A69587	intracellular alka
5	297.5	9.6	1398	2 T28159	pyrolysin (EC 3.4
6	286	9.2	806	2 A41341	microbial serine p
7	279.5	9.0	419	1 S25835	subtilisin (EC 3.4
8	276.5	8.9	799	1 C83753	subtilisin (EC 3.4
9	274.5	8.8	420	1 S23407	subtilisin (EC 3.4
10	272.0	8.7	580	2 S11850	serine proteinase
11	268.5	8.7	715	2 UC4508	alkaline serine pr
12	262	8.4	894	2 F69730	cell wall-associat
13	260.5	8.4	1345	2 T29090	surface layer-asso
14	255	8.3	378	2 A33933	high-alkaline seri
15	255	8.2	513	1 A35742	high-alkaline seri
16	250.5	8.1	380	2 A49778	aquafalysin (EC 3.4
17	250.5	8.1	601	2 UC4576	serine proteinase
18	250	8.1	382	2 I39780	subtilisin (EC 3.4
19	249.5	8.0	757	2 C84120	subtilisin-type pr
20	248	8.0	627	2 D75393	serine proteinase,
21	246.5	7.9	402	2 JU0332	alkaline proteinase
22	245.5	7.9	381	2 JH0778	subtilisin (EC 3.4
23	245	7.9	382	1 SUBSN	subtilisin (EC 3.4
24	243.5	7.8	381	1 SUBSN	subtilisin (EC 3.4
25	242.5	7.8	379	1 SUBSCL	subtilisin (EC 3.4
26	242	7.8	1331	2 A72647	probable surface l
27	241.5	7.8	381	1 SUBS	subtilisin (EC 3.4
28	239.5	7.7	401	2 I39974	serine proteinase
29	239.5	7.7	534	1 JS0173	alkaline proteinase

30	237.5	7.7	381	2 T01487	subtilisin (EC 3.4
31	235	7.6	519	2 S71451	halolysin R4 (EC 3
32	232	7.5	488	2 A11930	proteinase [import
33	231.5	7.5	1167	1 A35066	streptococcal C5a
34	230	7.4	384	2 UC4802	alkaline proteinase
35	223.5	7.2	409	1 S32905	serine proteinase
36	223	7.2	321	1 S27501	alkaline proteinase
37	219.5	7.1	613	2 S75976	hypothetical prote
38	218.5	7.0	272	2 A23624	subtilisin (EC 3.4
39	218.5	7.0	275	2 JC1085	subtilisin (EC 3.4
40	218.5	7.0	533	1 JU0146	serine proteinase
41	218.5	7.0	535	2 B82358	alkaline serine pr
42	218	7.0	361	2 G83756	subtilisin-type al
43	216.5	7.0	361	2 A48373	high-alkaline seri
44	216.5	7.0	525	2 G84406	halolysin [importe
45	216.5	7.0	645	1 SUBSMP	serine proteinase

ALIGNMENTS

RESULT 1  
T18279  
multidrug resistance transport protein - slime mold (Dictyostelium discoideum)  
C:Species: Dictyostelium discoideum  
C:Date: 15-Oct-1999 #sequence\_rev15ion 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T18279  
R:Shanisky, G.; Loomis, W.F.  
Submitted to the EMBL Data Library, June 1996  
A:Reference number: 218855  
A:Accession: T18279  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1743 <SMA>  
A:Cross-references: EMBL:U60086; NID:g1399914; PID:g1399915; PIDN:AA03331.1  
C:Genetics:  
A:Gene: tagc

Query Match 16.7%; Score 518.5; DB 2; Length 1743;  
Best local Similarity 24.2%; Pred. No. 2.4e-28;  
Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;

QY	88	HIXQFNGPILEETKQXL-----EXTGAKILDIYIPDAVIV---EYEGVAXSX 131	
DB	115	YIVQFDRINDETRQLEKFLIGTDIVLDQPYQSHIVYIPHDSFLVLTMQDSVLLS 174	
QY	132	XXXIEHVESVEPY-----LPYXIIDPOLFTKGASLVK--AXALDT--KQXN 174	
DB	175	KEWVSWIGFEPSKRIHLYNNEKISGLPYIT---LSDSTNSLIQKRENTLNSILKSYN 230	
QY	175	KEYQLRGIT-----EXIAQXXXSNDVXYITAPPEK 204	
DB	231	SKVKLTLLNOKKLSIYCNDESPSPCSLINSEKLVYOWISEQSEN--FLERSEKFO 287	
QY	205	VWMDVARGIV-----KADVAOSSYGLYGQGIYAVADTGLDTGR---NDS---- 246	
DB	288	TANRLSPKVFECTKDTLVNNDRDVDP-----LRGKOILSTADTGLDOSHCFSDSYPI 342	
QY	247	---SMHEAFRGKITALYALGRTNANDNTNGHTHVAAGSVLG-----NGXTNKGAPQA 296	
DB	343	PLNSVVALNLR-KVYVIYITSDSDSDKVDGCHTICGSAAGPESSVNISSPSGLATPA 401	
QY	297	NLVFQSIKMSXGGLGGL--PSNLOTLFQSAKSAIRHFNHNSGA-----AVNGAVTTDSRN 350	
DB	402	KIAF--FDLASGSSSLTPPSDLKQLXPLOYDAGARVHODSMGSVSEGTGVSISDTAS 458	
QY	351	VDDYVAKN-DMTILFPAAGNEKNGRTIS--APGTAKNAITGARENLR-----PSFGS 400	
DB	459	IDDFLETHDFITLRAQGN---NEQVLSLTOSTAKNVTYGAQOTTHEWYLDGQRYNYN 515	
QY	401	YADNT----- 405	
DB	516	YQSSVDINDELICDFDSRYCNVYTTAQCCLESNATYGLASCCPTLLRKSVIDAANTOPPLY 575	

```

Oy 406 ---NHVADSSNGPKRODIRITPDMAAGTXYLXKRSLA-----PDSSWMANHDSKAYM 457
Db 576 NENNCSFSSSGPTHDRCRM.PALVABGEYITSRSNGAATTDGCGGSL-PTNTALLA-I 633
Oy 458 GGTSMAPPIYVAGNVAOLREH-----FYVNRGITEPKPSILKAAALAGA----- 499
Db 634 SGTSMAITSFPAALAAATTTILROYLDVGYPTGSIYESNKLOPTGSLKRALMINNAQLINGTFQ 693
Oy 500 ---ADXXGIGPANGN-----OGMGRTTLDKSLNVAAYNESS----- 531
Db 694 LITSSSTIYFBNQYFENFAGASLYOGHGAIRMSNMLHYVNNNNNSNNNNKTSIDGITKFCGI 753
Oy 532 -----XLISTSOKATYXFT-----ATAGKPLK--ISLVSDAPASTT 565
Db 754 GGLDLRLLYVKPNQMKESLSTGQNTSYCFYKPPSSSSNSNGNINPRVAVATLVWTDPPSAG 813
Oy 566 ASVTLVVNDLIDIVT-----TAPNGTXYVGDXFXXPPXXXMMDGANNVENVFINXP 613
Db 814 AKFNLVNNLDLITMYIRYDNGSTIFYSNSOGSSFLG-----LAPQDVLNNEGIVHNPT 867
Oy 614 OSGTYTIEYQAVNVPVGPQKFS 635
Db 868 EPMYIRFVPVAGTNVMPGPNFS 889

```

## RESULT 2

multidrug resistance protein - slime mold (*Dictyostelium discoideum*)  
C:Species: *Dictyostelium discoideum*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T18267  
R:Shaulsky, G.; Kuspa, A.; Loomis, W.F.  
submitted to the EMBL Data Library, January 1995  
A:Description: An MDR transporter/serine protease gene is required for prestalk speciali  
A:Reference number: Z18850  
A:Accession: T18267  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1905 <SHA>  
A:Cross-references: EMBL:U20432; NID:q664839; PID:q664840; PIDN:AAA6212.1  
C:Genetics:  
A:Gene: tagB

Query Match	15.48;	Score 477.5;	DB 2;	Length 1905;
Best Local Similarity	24.58;	Pred. No. 2.1e-25;		
Matches 193; Conservative	98;	Mismatches 237;	Indels 259;	Gaps 31;

```

OY      88  HIXOPNGIIEETKOXLE-----XHGAKLIDYIPDYAIVEY-----EGDYXS 130
Db      172  YIVQKDRINIBETROALKFELIGITDITILKQBPKSHVHVIIPHDSFLVFWTKQSVLSS 231
OY      131  XXXXEHESEVEPY-----LPYXHI-----DPOLFETKGASL-----V 163
Db      232  SKEMTSMIGEHPKSHIHNIHENSIGYPIYIILSGTNSLIQWMENTLSILTYSKYV 221
OY      164  KAKALDTKQ-----XNKEVULRGIEIXIAQXXXS---NDVXYITAKPEKVM 206
Db      292  KTLINQKKLKLSIYCNDESPSSSSSSCSLIGSEKIIVYKMSISQSSMNIENREKQTA 351
OY      207  NDVANGIY-----KADVAQOSTGLYGQCIYAVADGTGLDTR---NDS----- 246
Db      352  NRIESTYFETGTRDKLVNNDRIDIP-----LEKGOIILSIDTGDIGDSCHEFSOSKYP1PF 406
OY      247  SMHEAPFGKLTALYALGRTNANNADNTGHHGHVAGSVLGNKXTH-----KGMAPQANL 298
Db      407  NOVNNENHKKVYT---YITHYHDENYIVMGHGHHVCGSAGAPEDSSWAISSFSGLADAKT1 463
OY      299  VFOSIMDSXGGLGGLPSNLIQFLFOAQSAGARIHTNSMGA-----AVNGAYTTSRANDDY 354
Db      464  AFYID-LSSGSSEPPRPEDYISQMYKPLVDAGKRVHGDMSGVSISLOGYIGGYSDAAGGIDAF 522
OY      355  VRK-NDMTILFPAAGNEKNGGTISAPGTAKNALTVGATENLRPSEGS-----YADNI- 405

```

```

Db      523  IYEYEFILRAQNN-ELFASLQAATAKNAITVGAEQTAHVYVSDALEYDESDNAN 584
      406  -----NHYAOF 411
07

```

```

Db      582  F0RPLLEFKKICNTYTTACCESEVSNVKGJLQCLCPASIKONASBFTTQPOFYENNNNGSF 641H
QY      412  SSRGPTKDGRIKPDVMAFGTYILSARSS-----LAPDSEFWANHDKRYAVMGTSNATP 465H
      ||||| |||||::||::||| ||||| : : : |||||
Db      642  SSKGPTHGRLKLPDIYAVAGEYITSAKRSNGEMSTQCGGSL- -PNANGMSISGTSNATP 699H
QY      466  IYAGNVAVOLREHF-----YKNRGITPSPULKAALI----- 496H
      ||::| :||::| :|||::| :|||:
Db      700  LATATFTILRQLYDGVPTGESVVEENKLLPTGSLIKALMINNQLNGTYEWSASSTNP 759H
QY      497  AGAAMXGGLGYPNGNOCWRVTLDKSLNAVYNNESS-----XLSISOKAT----- 546H
      :||:| :|||:| :|||:| :|||:| :|||:|
Db      760  SNAIFEOINGAALLQGWGALRKN--NMWLYKSSNPPTPPSSWIGIGGJGKNOKATEKED 816H
QY      541  -----YXFT-----ATAGK-P-LKISLWSDAPASTYASVTLVNLDDL- 576H
      ||||| :|||:| :|||:| :|||:| :|||:|
Db      817  SLSSGLKNSYCTTYPPSSSSSGSGGGGTTPRIYATLVWTDPPSISGAKFNLVNNLDLLL 876H
QY      577  -----VITAPN--GTAYVGNDFEXXPXKXNMNDGRNNENVEFNKPSGCTYITIEVOAYNP 628H
      ||||| :|||:| :|||:| :|||:| :|||:|
Db      877  NSDDDSITITIGNSGSLQPAKVAQP-----DTLNNEGIIITNPTKAMNYKFTIAGTNVP 931H
QY      629  VGPQXKFS 635
      :|||:|
Db      932  IGPQKFS 938

```

### RESULT 3

B56571 Intracellular alkaline serine proteinase aprX [imported] - *Bacillus halodurans* (strain C-125)  
 C:Species: *Bacillus halodurans*  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
 C:Accession: B83891  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hino, T.; Nakamura, K. *Microb. Drug Resist.* 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* strain C-125  
 A:Reference number: AB3650; MUID:20512582; PMID:11058132  
 A:Accession: B83891  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-444 <STO>  
 A:Cross-references: GB:AP001513; GB:BA000004; NID:J10174345; PIDN:BA805649.1; GSPDB:G000000000  
 A:Experimental source: strain C-125  
 A:Genetics:  
 A:Gene: aprX

Query Match	11.38;	Score 350.5;	DB 2;	Length 444;
Best Local Similarity	30.28;	Pred. No. 2.9e-17;		
Matches 114; Conservative	54;	Mismatches 127;	Indels 83;	Gaps 17;

```

0Y 1B3 E1IAOXXXXNDVXYITTAPEKVNANDVARG1YKA-DVAOSSYGLXGOGQ1YVAAPDLDLT 24 41
Db 100 BSLQEMLVCKD1RKY1YLNREHALD1D1AVESAQAPEY1RNET1YLGKDVT1YAV1D1G1 15 71
0Y 242 GRNDSSMEAEFRGKITALLY-ALGRTNANND1NGHGT1YAGSVLNGXTN---KGMAPOA 296
Db 158 -----YPHEDELEGRIKAVDEPNVNOREEPYDNDNGHTICAGDMAONGAS1SDBOYUGAPAEA 212
0Y 297 N1VFO5INDSKGLGGLPE5N1QTD1F5QAXSAGAR1HTN5MCAVANGAYTDSRNV----- 351
Db 213 N1VIGKVLNKQ-GMGSLE51M0GV-----EMCIQYNEHPDPIH1ISM5L 257
0Y 352 -----DDYVR-----KNDMT1LFAAGNEXPNGST1SAPCTAKAT1YVGTEN 399
Db 258 GQALPEYNEQEDBPWKR1YVEEAMNAG1YVCYAAAGSGCPDAQ1TASPGVSEY1YVGA1DD 317
0Y 394 LRP5FG5YADN1N1HVAQ5FSSRGKPTKDGRIKDPVMAPGT1Y1SANS1LAPD5SF-----WA 448

```





Db 246 NMSIGSSGESSLTIT---NAVDAYDKGVLIIAAGNSGPRGSGIGYALVNAVALE 302  
Qy 393 NLRPSFGSYDININVAOFSSRGPTKDG-----RIKPDVMACTIILARSILAPDSF 446  
Db 303 NTION-CTY-----RVADFSRGRHRTAGDVYICKGVEISAGAAVYST----- 346  
Qy 447 WANHDSKYAVMGSTSMATPIVAGNVAOL 474  
Db 347 W--FDGCVATISCTSMASPHAGLAKI 372

## RESULT 8

G83753  
subtilisin-type proteinase (EC 3.4.21.-) vpr precursor [similarity] - Bacillus haloduran  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: G83753  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hara  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: G83753  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-799 <STO>  
A:Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA04550.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: vpr  
C:Superfamily: microbial serine proteinase vpr: subtilisin homology  
C:Keywords: hydrolase, serine proteinase  
F:1-29/Domain: signal sequence #status predicted <STG>

Query Match 8.9%; Score 276.5; DB 2; Length 799;

Best Local Similarity 23.3%; Pred. No. 1; Le-11;

Matches 149; Conservative 61; Mismatches 198; Indels 227; Gaps 25;

Qy 78 LEIVPANNKLHIQENP-ILE-----ETKQXLEXTGAKILDVIPPAYIVYE 125  
Db 50 LEIV-----IVEIDPSIIIAKHOGKOSKNEKQARQSVIEQIDLVLP--SSTVTHE 100  
Qy 126 GDVAXSXKXIIIEHVESVEPYLPIYXIDPOLFTKGASXLYKAKALDTRQKKEVOLGIEIXI 185  
Db 101 YDFLESGFALE-----LPAHQIPISILGIDGVHAYVPIEVEVETDDEV----- 144  
Qy 186 AOXKXSNVXYITAKPEYK-VNNDVARGIVKADVAOSSYGLYOGGIVAVADTGLDTRN 244  
Db 145 -----VIEKDAYSPKMLDSAPFICANDAMEAGY--TGEGLTVALIDTGVDTYHP 191  
Qy 245 DSSMHEAFRCKITAYALORTNNANDT-----NGCHTVAGSVLNGXTNGAPQ 295  
Db 192 D--LVNAF-GDYKGMDFIDNNDPOETPPGCPRIETTHGTAVAGIVANGLI-KGVAPD 247  
Qy 296 ANLVFOSIMDSKGLGLPSNLQTL--FSQAXSAGARIHTNSGAANGAYTTDSRVND 353  
Db 248 ANLAIYFVL---GGGKRGSTAGVLAGIERVAVODGAOIMNLSGNTLNDPDEATIAL-D 302  
Qy 354 YVRKNDMTILFAANGEXPNGGTISAPGTAKNAITVGATENMLRPSFGSY-----ADNINH 407  
Db 303 WAMAEGVAVATSNNGSPNNMTVSGPSTSDAISVGAT--RLPYKRYKASVFTSDGIDY 359  
Qy 408 ----- 407  
Db 360 PSADIMGFPSEDELELDGETEYFAAGLAKPGDFEGVDEGKIALIVGEIPEVKAEN 419  
Qy 408 ----- 407  
Db 420 AKAGAVGATITNNVAGVQPIVGLAIPITIMLSNEDKLKRNLENGCONTVFSTIEPKL 479  
Qy 408 -----VAQFSRGPT-KDGRKPDVMACTIILARSILAPDSFANHDSKYAVMGSTSM 462  
Db 480 VGEIVADFSRGRVMTHTMIKIPVSAFVAIVSTIPTHODDPY-----GIGSKQGTSM 533

Qy 463 ATPVAGNVAOLREHFVKNRGITPRPSLLKAALTAGA-----DXGLGYPNGOGCRVTL 518  
Db 534 ASPRVAGAAALLLEAH-PWKGVY---DHVKAALMNTAENLVENGNRRPHTMGAG----- 584  
Qy 519 DKSILNAVYNESSXLSSTOKATY-XFTATGKPLK 552  
Db 585 --SIRIVDAIESETLVTSQSHSFGFTKRGKQVE 617

## RESULT 9

S23407  
subtilisin (EC 3.4.21.62) 1 precursor - Bacillus sp. (strain TA39)  
C:Species: Bacillus sp.  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 18-Jun-1999  
C:Accession: S23407  
R:Narinx, E.; Davail, S.; Feller, G.; Gerday, C.  
Biochim. Biophys. Acta 1131, 111-113, 1992  
A:Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarctic  
A:Reference number: S23407; MUID:92256481; PMID:1581352  
A:Accession: S23407  
A:Molecule type: DNA  
A:Residues: 1-420 <NAR>  
A:Cross-references: EMBL:X62369; NID:g40200; PIDN:CAA44227.1; PID:g40201  
C:Genetics:  
A:Gene: sub1  
C:Superfamily: subtilisin: subtilisin homology  
C:Keywords: extracellular protein: hydrolase; serine proteinase  
F:1-23/Domain: signal sequence #status predicted <STG>  
F:24-111/Domain: propeptide #status predicted <PRO>  
F:112-420/Product: microbial serine proteinase #status predicted <MAT>  
F:136-374/Domain: subtilisin homology <SBT>  
F:145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 8.8%; Score 274.5; DB 1; Length 420;

Best Local Similarity 28.3%; Pred. No. 6; 3e-12;

Matches 107; Conservative 56; Mismatches 142; Indels 73; Gaps 21;

Qy 168 LDRKQXN--KEYOLRGIEIXIAQXKXSNVXYITAKPE--YKVV---NDVARGIV-----KA 216  
Db 73 MNEKOFALKKKNKNIIVKEVPEL---EIAATDKPEALYNMAASQSTPMGKIAYNNS 128  
Qy 217 DVAGSSYGLYGQGOIYAAVADTGLDTRNDSMHEAFKGTAL--VALGRT---NNANDT 271  
Db 129 SIQGTG---GGGINIAVLDTGVNTN-----HPDLRNNVEQCKDFVQTYTNNSCDR 179  
Qy 272 NGCHTVAGSVLNGXTNGAPQ---GMAPQANLVFOSIM--DSXGLSGISPNLQTLFSQAXS 326  
Db 180 QGCHTVAGSALADGDTGNGYVAPADMLAKYVLDDSCGADDAIAAITHAGDQATA 239  
Qy 327 AGARIRTN-SWGAANGAYTTDSRVNDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNA 385  
Db 240 LNRKVIYNNMSLSSGESSSLITNAVN---YSYNGVLIILAAAGNSGYPQOSIGYPAALVNA 296  
Qy 386 ITVGATENLRPSFGSYADININVAOFSSRGPT-KDG-----RIKPDVMACTIILARS 439  
Db 297 VAAVALEEN-KVENGYV-----RVADFSRGRYSWTDGYALOKQDEISAPGAIYST--- 347  
Qy 440 LAPDSSFWANHDSKYAVMGSTSMATPIVAGNVAOLREHFVKNRGITPRPSL-----L 491  
Db 348 -----W--FDGCVATISCTSMASPHAGLAKIYAVQPSANVDVREGELQYRAYENDI 388  
Qy 492 KAALINGAAD---XGIGY 506  
Db 399 LSGYVAGYGDDEAFASGPGF 416

## RESULT 10

S11890  
serine proteinase (EC 3.4.21.-) precursor, extracellular - Xanthomonas campestris pv.  
N:Alternace names: subtilisin-related proteinase  
C:Species: Xanthomonas campestris pv. campestris  
C:Date: 21-Nov-1993 #sequence\_revision 07-Feb-1997 #text\_change 03-Dec-1999

[illegible]

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-715 <TSU>  
A:Cross-references: DDBJ:D38600; MID:q1536787; PIDN:BAI8912.1; PID:dt1019647; PID:921  
A:Experimental source: strain O-7  
C:Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-se  
C:Genetics:  
A:Gene: aprI  
C:Superfamily: subtilisin homology  
C:Keywords: hydrolase  
F:1-40/Domain: signal sequence #status predicted <Sig>  
F:41-150/Domain: amino-terminal propeptide #status predicted <Amp>  
F:151-496/Product: alkaline serine protease I #status predicted <Mat>  
F:181-452/Domain: subtilisin homology <SBT>  
F:497-715/Domain: carboxyl-terminal propeptide #status predicted <CPT>  
F:239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 8.7%; Score 268.5; DB 2; Length 715;  
Best Local Similarity 25.8%; Pred. No. 3,5e-11;  
Matches 142; Conservative % 51; Mismatches 201; Indels 157; Gaps 28;

QY 169 DTKQKNKEVOLRGEXINQXXXNDVYXIT-----AKREKRVMD-----VANGIV 214  
DB 117 DKLSAKAPQAPMNVVA-----SGNVEYIETDQMKRPATPNDPRYNDQMHYAAAGT- 171  
QY 215 KADVAOSSYGLYGQGOIYAVADTG---LDTGRDSSMHEAFRGKITALYALGFTNAND 270  
DB 172 NAPAAMDK-ATGCGVVVAVALDTGYRPHLDANILPGDYIMISMTFVAANDGARDNARD 229  
QY 271 -----TNG-----HGHVAG---SVLGNXTYKGMAPQANLYEQS 302  
DB 230 PGDAVTRBECGTDSGQVPVPRADDSSMWHGTHVAGTVAALVNNEGVAGAAVYDAKKVVR 289  
QY 303 IMDSXGGLGGLPSNIQTLFQOAXSAGARIHTNSGAAV-----NGAYTDSRVVDYV 355  
DB 290 VL---GKCGGLSLDADIDIMASGSDRVPANANPAVYINNSLGCGACSATGQNALNQ 346  
QY 356 RKNDMTLFLAAGNEXPNGCTISAPCTAKNAITVQAT--ENLRPFGSTADNINHVAQFSS 413  
DB 347 RNNGVIVYIAAGNDNDSNANYN-PGNCNGVYVNAVSVGRDGSRAVYSNNGANI----- 397  
QY 414 RQPTDGRITKPPVMAPGTXILSARSILAPDSSFPANMHS-----KYAVMGSTMATP 465  
DB 398 -----DVAAPG---GAQSFADDPGGLISTHNSGSGAASNSYHSGQTSMAAP 442  
QY 466 IVAGNVAAQLR-----EHFVKN--RGITRPSILKALTI--AGADXGLG-----Y 506  
DB 443 HVAGVAAALIKQAKPSATPDEVEETILKNTTRSFAGSCSNCGTGVDAALAAVNEALGADVTP 502  
QY 507 PNGNGQWGRVTLDSKLVNAAVYNESXLISTSGKATYXFPATGKPLKLSLVMSDAPASTTA 566  
DB 503 PFGN-----TLED--GYAKTGLSGAAGSNQ--FFTEVPYAGK-----TNV 538  
QY 567 SVTL--VNDLDLVITAPNGTXYYGNDPEXDXPKXNMGR---NNVENVFNKQSGTYT 619  
DB 539 PFTMSGGTGDADLVYK-----LGSQ---PTSSSDCRPEYGNAEVCSPDADQAGTYH 588  
QY 620 IEVQAYNPVG 630  
DB 589 VMINGKAYSG 599

RESULT 12  
P69730  
Cell wall-associated protein precursor wpra - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence,revision 05-Dec-1997 #text\_change 15-Oct-1999  
C:Accession: P69730  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A:Enrich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabeet, C.; Ferrari,  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal



leach, J.: Harwood, G.R.; Hensut, A.; Hilbert, H.; Holaspee, S.; Hosono, S.; Hullio, M.  
Koeffer, P.; Konigstein, G.; Krogh, S.; Kumano, T.; Kunita, K.; Lapidus, A.; Larinios,  
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauec  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Paro, V.; Poln, T.M.; Poterallio,  
Rieger, M.; Rivolta, C.; Rochoa, E.; Rochoa, B.; Rose, S.; Sadate, Y.; Sato, I.; Scanon  
A.; Authors: Schleib, S.; Schoeber, R.; Scoffone, F.; Seidguth, J.; Sekowsky, A.; Serco  
krenchl, M.; Tanakaoshi, A.; Tanaka, T.; Teyssie, P.; Tsunamoto, K.; Yamamoto, A.;  
T. Winters, P.; Witat, A.; Yamamoto, H.; Yamane, K.; Yamoto, K.; Yata, K.; Yoshida,  
A.; Authors: Yoshiyama, H.F.; Zumbstein, E.; Yoshiyama, H.; Danchin, A.; Bacillus subtilis.  
A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A:Accession number: A69580; MWID:98044033; PMID:9384377  
A:Reception: F69730  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1894 <RUN>  
A:Cross-references: GB:I299109; GB:AL009126; MID:g2633260; PIDN:CAB12917.1; PID:e1183079.  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: wpiA  
C:Superfamily: subtilisin homology  
C:Superfamily: subtilisin homology <SBT>  
I:43-664/Domain: subtilisin homology <SBT>

```

Query Match      8.4%; Score 262; DB 2; Length 894;
Best Local Similarity 23.9%; Pred. No. 1.4e-10;
Matches 134; Conservative 92; Mismatches 211; Indels 124; Gaps 28;

OY 50 ESKRQXGQAAAF--LVESENVLAKGLKKKLETPVA-----NKKLHXQFN-----G 94
D 245 FDDVSENGASSSKRYVEKQKAINRLYDKALDSVSEFLKEETKQADRLNKLQGLQGTAG 304
OY 95 PILEE-----TKQXLEKTXGAKILIDYIDPAYIV---EYEG-----DVSXXXXX 134
D 305 ALLENNIAAKSEVQTT--KVIFKVDNKSLSSVHNEMKGPSASASQSKDIDSNVKKRAKL 362
OY 135 IEHVESVEPYLPYXXIDPOLFTKGASXLVKAHALDTKQXNEVQLRGIEIXIAQ--XXXXND 193
D 363 FDNLYSFE--LPDEKQONGAYTASAKRYKSAATLTKSSNVEF---AEPQEKYSKL 416
OY 194 VXYITAPKPEKVNVDVARGIV--ADVA--OSSYGLGQGO---TVAAVADGLDTRGRDSS 247
D 417 IQY---PYQPLKLNNEENGSGVKADYKKEPANTLLSKRLKMLDILAVVDYGTSTLAD-- 471
OY 248 MHEAFRCKITALA---LGRTNANDNTNGCHGVAGSVLG---NGXINKMAQANLVFO 301
D 472 ---LKKKVTVDLGHNFVGRNNAMDDQGHGVAGIILAAOSDMSGYSTGLMNAKAITPY 527
OY 302 SIMDSKGGGLGCLPSNLQTLFQASXSAGARIHTNSNGAANVANGAYTSDSRNVDVYRKNDMT 361
D 528 KVLDSAC--SGDEQJQALGIKYAADKAGKAVILSLG---GGSYRVLEFFALKYAAKKNYL 581
OY 362 ILFPAENXENPGTGIIAPGTAKNATIVGATEMLRPSFSGSYADNTNINHAQSSRCPTDGR 421
D 582 IAAASGNDGEN--ALSYPASSKRYMSVGT-----NRMDMTADPSNGKGL--- 625
OY 422 IKRPVMAPEGXIIILASRSLAPDSSFANHDSKYAANGSTNMAPTIVGANAOLREHFVKN 481
D 626 ---DISAPGSDI---PSLVPNGN-----YITMSTSNATPYAAAAAGLFLAQNPKL 670
OY 482 RGIIRKSLKAA--LIIAAGDXG-----LGYP-----NGQOGKRWTLDKSLN 523
D 671 KRTEVEMLKKTADDISFESVDGGEELDYDGPDIETKPPGVDMHSGTGRVNMKAVS 730
OY 524 VA---YVNESSXLTSTQKAT 540
D 731 AADQLQKVNKLESTQTAIVAGS 751

```

RESULT 13  
T29090  
surface layer-associated STABLE proteinase - Staphylothermus marinus  
N: Alternate names: hyperthermostable proteinase  
C: Species: Staphylothermus marinus  
C: Date: 02-Sep-2000 #sequence\_rev1stion 02-Sep-2000 #text\_change 02-Sep-2000

C:Accession: T29090  
C:Myrv, J.; Lups, A.; Kellermann, J.; Eckerskorn, C.; Bammelster, W.; Peters, J.  
Curr. Biol. 6, 739-749, 1996  
A:Title: A hyperthermostable protease of the subtilisin family bound to the surface  
A:Reference number: Z20539; PMID:96385442; PMID:8793300  
A:Accession: T29090  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1345 <MANY>  
A:Cross-references: EMBL:U07968; NID:g1374755; PID:g1374756; PIDN:AAB02333.1  
A:Experimental source: strain F1  
C:Function:  
A:Description: probably serves an exodigestive function related to the organism's end  
A:Note: stoichiometric S-layer component

[illegible]

RESULT 14  
A33973  
high-alkaline serine proteinase (EC 3.4.21.-) Yab precursor - *Bacillus* sp. (strain Yab)  
N:Alternate names: alkaline elastase; subtilisin Yab  
C:Species: *Bacillus* sp.  
C:Date: 23-Mar-1990 <sequence\_revision 23-Mar-1990 <text\_change 22-Jun-1999  
C:Accession: A33973; B33973  
J:Kaneko, R.; Koyama, N.; Tsai, Y.C.; Juang, R.Y.; Yoda, K.; Yamasaki, M.  
A:Title: Molecular cloning of the structural gene for alkaline elastase Yab, a new strain of *Bacillus* sp.  
A:Reference number: A33973; MUID:89359181; PMID:2670913  
A:Accession: A33973  
A:Molecule type: DNA  
A:Residues: 1-378 <KAN>  
A:Cross-references: GB:J28537; NID:g341960; PIDN:AAA8734.1; PID:g758668  
A:Note: Parts of this sequence, including the amino end of the mature protein, were determined from the complementary DNA sequence.  
A:Accession: B33973  
A:Molecule type: protein  
A:Residues: 111-164;326-355 <KA2>  
C:Genetics:  
A:Gene: ale  
A:Start codon: TTG  
C:Superfamily: subtilisin; subtilisin homology  
C:Keywords: extracellular protein; hydrolase; serine proteinase; zymogen  
F:1-27/Domain: signal sequence #status predicted <StG>  
F:28-110/Domain: activation peptide #status predicted <ACP>  
F:111-378/Product: alkaline elastase #status predicted <MA7>  
F:132-338/Domain: subtilisin homology <SB7>  
F:141,171,324/Active site: Asp, His, Ser #status predicted



GenCore version 5.1.4-P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:28:13 ; Search time 10.0078 Seconds  
(without alignments)  
2652.411 Million cell updates/sec

Title: US-09-920-954-2

Perfect score: 3102  
Sequence: 1 MKKKKKVFLSVLSAAAILST.....EVOAYNPVGPOKFSALIVN 640

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Swissprot\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518.5	16.7	1743	1 TAGC_DICDI	Q23868 dictyostell
2	477.5	15.4	1905	1 TAGB_DICDI	P54683 dictyostell
3	299.5	9.7	1398	1 PLS_PYRPU	P72166 pyrococcus
4	286.5	9.2	806	1 SUBV_BACSU	P29141 bacillus su
5	274.5	8.8	420	1 SUBV_BACSU	P28842 bacillus su
6	270	8.7	580	1 PROA_XANCP	P23314 xanthomonas
7	262	8.4	894	1 WPRX_BACSU	P54423 bacillus su
8	257.5	8.3	378	1 ELXA_BACSP	P20722 bacillus su
9	255	8.2	513	1 NQLL_THRXO	P08394 thermus aqu
10	250.5	8.1	380	1 ELXA_BACAO	P27693 bacillus al
11	246.5	7.9	380	1 ELXA_BACCS	P41362 bacillus cl
12	245.5	7.9	402	1 ALP_CEPAC	P29118 cephalospor
13	245.5	7.9	381	1 SUBV_BACNA	P35835 bacillus su
14	245	7.9	382	1 SUBV_BACNA	P00782 bacillus su
15	242.5	7.8	379	1 SUBV_BACLI	P00780 bacillus su
16	241.5	7.8	381	1 SUBV_BACSA	P00783 bacillus su
17	239.5	7.7	381	1 SUBV_BACSU	P04189 bacillus su
18	239.5	7.7	401	1 THES_BACSP	P45670 bacillus su
19	239.5	7.7	534	1 PROA_VIBAL	P16588 vibrio algi
20	239	7.7	269	1 SUBS_BACLE	P29600 bacillus le
21	238	7.7	269	1 PRTM_BACSP	O99405 bacillus le
22	237.5	7.7	381	1 SUBV_BACST	P29142 bacillus st
23	234.5	7.6	1181	1 SCAS2_STRPY	P58099 streptococc
24	232	7.5	269	1 SUBB_BACLE	P29599 bacillus le
25	231.5	7.5	1167	1 SCAL_STRPY	P15926 streptococc
26	223.5	7.2	404	1 ALP_TRIHA	O99778 magnetotric
27	223.5	7.2	409	1 ALP_TRIHA	O99778 magnetotric
28	223	7.2	321	1 SUPR_BACCS	O03420 triphodema
29	218.5	7.0	275	1 SUPR_BACPU	P07518 bacillus cl
30	218.5	7.0	533	1 PERC_ASPPG	P33258 aspergillus
31	218	7.0	361	1 ELXA_BACND	P41363 bacillus ha
32	216.5	7.0	642	1 SUBE_BACSU	P16396 bacillus su
33	214.5	6.9	326	1 ISP_PAEPO	P29139 paenibacilli

34	214	6.9	293	1 PRTT_TRIAL	P20015 tritirachlu
35	212.5	6.9	274	1 SUBD_BACLI	P00781 bacillus li
36	212	6.8	603	1 BPRV_BACNO	P42779 bacteroides
37	211.5	6.8	1052	1 MSIP_CRIGR	O92288 cricetulus
38	211.5	6.8	1052	1 MSIP_MOUSE	O92282 mus musculu
39	211.5	6.8	1052	1 MSIP_RAT	O92283 ratulus norv
40	211	6.8	530	1 HLY_HAL17	P29143 halophillic
41	210.5	6.8	1052	1 MSIP_HUMAN	O14703 homo sapien
42	210	6.8	422	1 TKSU_PYRKO	P58502 pyrococcus
43	210	6.8	595	1 BPRV_BACNO	P42780 bacteroides
44	210	6.8	1433	1 SUBE_BACSU	P16397 bacillus su
45	209	6.7	408	1 SEPR_THESR	P00146 thermus sp.

## ALIGNMENTS

RESULT 1	TAGC_DICDI	STANDARD:	PRT: 1743 AA.
ID	TAGC_DICDI		
AC	Q23868:		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Prestalk-specific protein tagc precursor (EC 3.4.21.-).		
GN	TAGC		
OS	Dictyostellium discoideum (Slime mold).		
OC	Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.		
OX	NCBI_TaxID=44689;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-AX4;		
RX	MEDLINE-97140317; PubMed-8986798;		
RA	Shaulsky G., Escalante R., Loomis W.F.;		
RT	"Developmental signal transduction pathways uncovered by genetic suppressors."		
RT	Proc. Natl. Acad. Sci. U.S.A. 93:15260-15265(1996).		
RL	PROC. NATL. ACAD. SCI. U.S.A. 93:15260-15265(1996).		
CC	-1- FUNCTION: INTERCELLULAR COMMUNICATION VIA TAGC MAY MEDIATE INTEGRATION OF CELLULAR DIFFERENTIATION WITH MORPHOGENESIS (BY SIMILARITY).		
CC	-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO PEPTIDASE FAMILY S8.		
CC	-1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.		
CC	-1- SIMILARITY: STRONG, TO TAGB.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	EMBL: 060086; AAB0331.1; -		
CC	DictyDb: DD02060; tagc.		
DR	InterPro: IPR003593; AAA_ATPase.		
DR	InterPro: IPR003439; ABC_transport.		
DR	InterPro: IPR001140; ABCtransportTM.		
DR	InterPro: IPR000209; Peptidase-S8.		
DR	Pfam: PF00005; ABC_tran; 1.		
DR	Pfam: PF00664; ABC_membrane; 1.		
DR	SMART: SM00382; AAA; 1.		
DR	SMART: SM00723; SUBTILISIN.		
DR	PROSITE: PS00136; SUBTILASE_ASP; FALSE_NEG.		
DR	PROSITE: PS00137; SUBTILASE_HIS; 1.		
DR	PROSITE: PS00138; SUBTILASE_SER; FALSE_NEG.		
DR	PROSITE: PS00211; ABC_TRANSPORTER; 1.		
KW	Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane; Signal.		
KM	SIGNAL.		
FT	CHAIN 1 27		POTENTIAL.
FT	CHAIN 28 1743		PRESTALK-SPECIFIC PROTEIN TAGC.
FT	DOMAIN 316 642		PROTEASE.

```

FT DOMAIN ? 1743 ABC TRANSPORTER.
FT TRANSMEM 962 982 POTENTIAL.
FT TRANSMEM 1027 1047 POTENTIAL.
FT TRANSMEM 1072 1092 POTENTIAL.
FT TRANSMEM 1157 1177 POTENTIAL.
FT TRANSMEM 1260 1280 POTENTIAL.
FT TRANSMEM 1288 1308 POTENTIAL.
FT ACT_SITE 325 325 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 637 637 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NP_BIND 1485 1492 ATP (POTENTIAL).
FT DOMAIN 42 46 POLY-ASN.
FT DOMAIN 94 103 POLY-ASN.
FT DOMAIN 643 646 POLY-ALA.
FT DOMAIN 733 741 POLY-ASN.
FT DOMAIN 786 792 POLY-SER.
FT DOMAIN 1337 1340 POLY-GLU.
FT DOMAIN 1346 1352 POLY-GLU.
FT DOMAIN 1353 1357 POLY-ASN.
FT DOMAIN 1358 1364 POLY-ASP.
FT DOMAIN 1381 1386 POLY-ASN.
FT DOMAIN 1707 1729 POLY-ASN.
FT CARBOHD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 689 689 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 735 735 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 741 741 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 832 832 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 887 887 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 1251 1251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 1385 1385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 1386 1386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 1454 1454 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 1704 1704 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1743 AA; 194145 MW; 12DB363E2F29839 CRC64;

```

Query Match Best Local Similarity 24.2%; Score 518.5; DB 1; Length 1743;  
Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;

```

QY 88 HIXQNGPILEETKQXL-----EXTGAKILDYIPDYIVY---EYEGDVXSK 131
DB 115 YIVGKRDINDETROQLKEFLIGTDIVLDEQYQSHIYIHPDSFLVLMTOEQSVLSS 174
QY 132 XXXIHEVSEVEPY-----LPIYXIDPOLFTKGSXLYK--AXALDT--KQXN 174
DB 175 KEWAWISEEPEPSNKHILNYEKSIGLPVYII---LSDSTNSLIQRMENLTNSILKSYN 230
QY 175 REVOLRGJ-----EXINQXXXNDVXYITPAKPEYK 204
DB 231 SKVKLTLLNQKKLKIIVYCNDESPSPSCSLINSEKLYIOWISEQSESN---FIENSEKFQ 287
QY 205 VMNDVARGIV-----KADYAGSSYGLYGQIVAAVADTGLDGR--NDS---- 246
DB 288 TANRLSPRVFEGTKDTLVNNDVRDIP-----LRGKQILSIADTGLDGHCFSDSKYPI 342
QY 247 ---SMHEAFRGKITALLVLRNNANNDINGHCTHYAGSYLG-----NGXTNKKMADPA 296
DB 343 PLNSVNLNHR-KVYVYITTTSDSDSKVDGHTICGSAAGPEPSSVIVSSISFGLADDA 401
QY 297 NLVPSIMDSXGGLGL--PSNLQTLFQOAXSAGARHITNSMGA-----AVNCAVYTTDSRN 350
DB 402 KIAF---FDLASGSSSLTPPSDLKOLYPLDAGARVHCDWSGVSVEGYTSYSDPAS 458
QY 351 VDDVYRKN-DMTILFAAGNEXPNGTIS--APGTAKNAITVATENLR-----PSFGS 400
DB 459 IDDFLFTHPDILLFAAQN---NEGYLSLTLQSTAKNAVITYGAHGTIHENITLIDEPNIN 515
QY 401 YADNI----- 405

```

```

DB 516 YOSSVDINQELICDFDSRYCNYTAQCCLESNATTGLASCPPTLRKSYIDAANTQPLLY 575
QY 406 --NHVAFSSRGPTKDKRKPDMVAPGTXXLSARSSLA-----PDSFEMAHDSKYAYM 457
DB 576 NENNICSSRSKPGTHDGMKPAALVAPGEYIISARSNGANTTDQCGDSL-ENTNALLA-I 633
QY 458 GGTSMATPIVAGNVAQLREH-----FYKNRGITPKPSLKAALIAA----- 499
DB 634 SGTSMATSEFAAATTLIKQYLVLDGYYPGSIYVESKRLQPTGSLKALMINNAQLNGTFQ 693
QY 500 --ADGGLGYNGN-----OGMGRVTLDKSLNVAAYNESS----- 531
DB 694 LITSSITPPSNQVFNENFAGASLVQGWCAIRSMNLHVYNNNNNNNNKTSDDGTTKFGCI 753
QY 532 -----XLTSGKATYXFT-----ATAGPLK--ISLWSDAPASTT 565
DB 754 GGLDLRLVKNQMKRESLSTGNTSYCTFYKSSSSSSNGSNIPRYVALTWTPDPSPVAG 813
QY 566 ASVTLVNDLDLYI-----TAPNGTYVGNDFXPPXKXNWDGRNVENVFTIXP 613
DB 814 AKFNILVNLDMITMYRRDNGSTIEFYSNQGSSEFLG-----LAPQDTLVNVEGIVHNPT 867
QY 614 QSGTYTIEVQAVNPVGPQXES 635
DB 868 EPMYRFVAGTVNVPMPQNFES 889

```

## RESULT 2

ID TAGB\_DICDI STANDARD: PRT: 1905 AA.

AC P34683:

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DT Prestalk-specific protein tagb precursor (EC 3.4.21.-).

GN TAGB.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

OX NCBI\_TaxID=44689;

RP [1]

RC STRAIN=AX4;

RE MEDLINE=95262903; PubMed=7744252;

RA Shaulsky G., Kuspa A., Loomis W.F.;

RT "A multidrug resistance transporter/serine protease gene is required for prestalk specialization in dictyostelium.";

RL Genes Dev. 9:1111-1122(1995).

CC -1- FUNCTION: INTERCELLULAR COMMUNICATION VIA TAGB MAY MEDIATE INTEGRATION OF CELLULAR DIFFERENTIATION WITH MORPHOGENESIS.

CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO PEPTIDASE FAMILY S8.

CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.

CC -1- SIMILARITY: STRONG, TO TAGC.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: U20432; AAA6212.1; .

DR MEROPS: S08.0PW; .

DR Dictydb: DD02059; tagb.

DR InterPro: IPR003593; AAA\_ATPase.

DR InterPro: IPR003439; ABC\_transporter.

DR InterPro: IPR001140; ABCtransportTM.

DR InterPro: IPR000209; peptidase\_S8.

DR Pfam: PF00005; ABC\_tran.1.

DR Pfam: PF00664; ABC\_membrane.1.

DR PRINTS: PR00723; SUBTILISIN.

```

DR ProDom: PD000006; ABC-transport; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00136; SUBTILASE-ASP; FALSE-NEG.
DR PROSITE: PS00137; SUBTILASE-HIS; 1.
DR PROSITE: PS00138; SUBTILASE-SER; 1.
DR PROSITE: PS00211; ABC-TRANSPORTER; 1.
KW Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane;
KW Signal.
FT SIGNAL. 1 31 POTENTIAL.
FT CHAIN 32 1905 PRESTALK-SPECIFIC PROTEIN TAGB.
FT DOMAIN 378 700 PROTEASE.
FT DOMAIN 1011 1031 ABC-TRANSPORTER.
FT TRANSMEM 1076 1096 POTENTIAL.
FT TRANSMEM 1121 1141 POTENTIAL.
FT TRANSMEM 1210 1230 POTENTIAL.
FT TRANSMEM 1309 1329 POTENTIAL.
FT TRANSMEM 1332 1352 POTENTIAL.
FT ACT_SITE 387 387 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 432 432 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 695 695 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NP_BIND 1553 1560 ATP (POTENTIAL).
FT DOMAIN 63 67 POLY-GLN.
FT DOMAIN 95 104 POLY-ASN.
FT DOMAIN 107 134 POLY-ASN.
FT DOMAIN 311 321 POLY-SER.
FT DOMAIN 833 837 POLY-SER.
FT DOMAIN 838 844 POLY-GLY.
FT DOMAIN 871 876 POLY-ILE.
FT DOMAIN 1012 1015 POLY-ILE.
FT DOMAIN 1386 1389 POLY-GLU.
FT DOMAIN 1398 1404 POLY-GLY.
FT DOMAIN 1445 1450 POLY-ASN.
FT DOMAIN 1765 1778 POLY-ASN.
FT DOMAIN 1782 1785 POLY-SER.
FT DOMAIN 1807 1812 POLY-PRO.
FT DOMAIN 1813 1850 POLY-GLN.
FT DOMAIN 1813 1878 POLY-PRO.
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 823 823 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1172 1172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1522 1522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1658 1658 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1905 AA; 212518 MW; B8E23FAB8B9AE13C CRC64;

```

```

Query Match 15.4%; Score 477.5; DB 1; Length 1905;
Best Local Similarity 24.5%; Pred. No. 4,9e-27;
Matches 193; Conservative 98; Mismatches 237; Indels 259; Gaps 31;

```

```

OY 88 HIXGNPILFETKXLE---XTGAKLIDYIPDYATIVY-----EGDYXS 130
DB 172 YIVGFKRINDETRQKLEFLIGITITLTKQPKRSHVHPIHDSFLVPMTRKESVLIS 231
OY 131 XXXXIEHVESVEPY-----LPXYXI-----DPOLFTKGASXL-----V 163
DB 232 SKEMISNIGHEPSNKHILNHEKSIQYPIVYIILSGSTNSLIQRMWNTJLSITLVSNSKV 291
OY 164 KAXALDTKQ-----XNKEVOJRGIEIXIAQXXXS--NDVYXITAKPEYVM 206
DB 292 KLTILNKKLKSIVYCNDESPSSSSSCSLIGSEKIVYMWIMISEQSESNVIRSEKLOTA 351
OY 207 NDVARGIV-----KADVAOSSYGLYGOGQIVAVADGILDTGR---NDS----- 246
DB 352 NRLSPVYIFGTKDKLVNNDRIDIP-----LRGKGQILSIADTGLDGSCHCFSDSKYPIPF 406
OY 247 -SMHEAFRGITLALVLRGTNNANDTNGHGTAVGSLGNGXIN-----GMPAPOINTL 298
DB 407 NOVMEHNRKRYVT---YITYHNEDEVVNGHGHVCGSAGAPPEQSSMAISSFGLATDKKI 463
OY 299 VFGSINMSXGGLGGLPSNLQTLFSAQXASGARIRHTNSWGA-----AANGAVTTDSKNVDY 354

```

```

DB 464 AFYD-LSSGSSEPTPEPEDYSQMYKRLYDAGARVGDMSKSVSLGGYGGSDDAGGIDAF 522
OY 355 VRK-NDMTILFAAGNEXPAGGTISAPGTAKNATVGTENTLPPSGS-----YADNI- 405
DB 523 LYEPERSILRAAGNN-ELFASILAQATAKNATVGAEGQAHVNVSDALEYDFSDNAN 581
OY 406 -----NHYAOF 411
DB 582 FQRCLEFDKKYCYNTTAKCCSEVSNVWGLQCCPASTIKONASDFTTQPOFENNMGSF 641
OY 412 SSRGPTDGRKRDVWABCTYILSARRS-----LAPDSFPAHNDSKYAAWGTSMATP 465
DB 642 SSKGPTHDGLKFDVABPEYITTSARRSGENSDIDCCDGL-PAHNGCLMSISGTSMAITP 699
OY 466 IVAGNVAQLREHF-----YKNGCITPKPSLLKALIT----- 496
DB 700 LATVAATTLNQYLVDGYPPTGESVEENKLEPTGSLKALMINNAOLLNGTYFWSASSTNP 759
OY 497 AGAADXGLGYPNGDGMGRVTLDSLVNAVYNESS-----XLSQOKAT- 540
DB 760 SMAIFEDINGANLIQWGLARMN--NMLYKSSNPTPPSRWIGIGIGLCKNOKATEWKED 816
OY 541 -----YXFT-----ATACKP-LKISLVSDAPASTAVTVNOLDL- 576
DB 817 SLSSGLKNSYCFYKPSSSSSSGGGGTPRIVATLVMTDPPSYSAKPNLVNNDLLLL 876
OY 577 -----VITAPN--GRXYVGNDFXKPPXXMMDGRNVEYFNXQSGTYITEVOAVNP 628
DB 877 NSDDDSITITIGNSGSLQPKAGVAP-----DTLNVVEGIILNPKANKYKFTTAGTNVP 931
OY 629 VGPQXFS 635
DB 932 IGPQKFS 938

```

RESULT 3

```

PLS_PYRFU STANDARD; PRT: 1398 AA.
ID PLS_PYRFU
AC P72186;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pyrolysin precursor (EC 3.4.21.-).
GN PLS OR PF0287.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN 111
RP SEQUENCE FROM N.A., SEQUENCE OF 150-184, AND CHARACTERIZATION.
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEHLIHE=9635370; PubMed=8702780;
RA Voorhorst W.G.B., Eggen R.I.L., Geerling A.C.M., Platteau C.,
RA Sleszen R.J., de Vos W.M.;
RT "Isolation and characterization of the hyperthermostable serine
RT protease, pyrolysin, and its gene from the hyperthermophilic archaeon
RT Pyrococcus furiosus."
RL J. Biol. Chem. 271:20426-20431(1996).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN 131
RP CHARACTERIZATION, AND 3D-STRUCTURE MODELING.
RX MEHLIHE=21079021; PubMed=11210516;
RA de Vos W.M., Voorhorst W.G.B., Dijkstraaf M., Kluskens L.D.,
RA van der Oost J., Sleszen R.J.;
RT "Purification, characterization, and molecular modelling of pyrolysin
RT and other extracellular thermostable serine proteases from
RT hyperthermophilic microorganisms."

```



RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RX MEDLINE-95020537; PubMed-7934828;  
 RA Glaser P., Kunst F., Arnaut M., Coudart M.P., Gonzales W.,  
 RA Hullo M.F., Ionescu M., Ludochinsky B., Marcelino L., Moszer I.,  
 RA Pressecan E., Santana M., Schneider E., Schweizer J., Vertes A.,  
 RA Rapoport G., Danchin A.;  
 RT "Bacillus subtilis genome project: cloning and sequencing of the 97  
 kb region from 325 degrees to 333 degrees.";  
 RL Mol. Microbiol. 10:371-384(1993).  
 RM [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RX MEDLINE-98044033; PubMed-9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Betero M.G., Bessieres P., Bolojin A., Borcherdt S.,  
 RA Borries R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capiano V., Carter N.M.,  
 RA Chol S.K., Codani J.J., Conerton I.F., Cumings N.J., Daniel R.A.,  
 RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabre C., Ferrati E., Foulger D.,  
 RA Frits C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,  
 RA Gilm S.I., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
 RA Guiseppi G., Guy B.J., Haga K., Hache J., Harwood C.R., Hentat A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Ilaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Kiebert-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Paro V., Pohl T.M., Portecelle D., Portwollik S., Prescott A.M.,  
 RA Pressecan E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeder R., Scofield F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takekura K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terstra P., Tononari K.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vanlier F., Vassaretti A.,  
 RA Viari A., Wambolt R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 subtilis".  
 RL Nature 390:249-256(1997).  
 CC - FUNCTION: NOT REQUIRED FOR GROWTH OR SPOULATION.  
 CC - SUBCELLULAR LOCATION: Secreted.  
 CC - PFM: PROBABLY UNDERGOES C-TERMINAL PROCESSING OR PROTEOLYSIS.  
 CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isp-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M76590: AAA22881.1;  
 DR EMBL: X73124: CAA51601.1;  
 DR EMBL: Z99123: CAA15835.1;  
 DR FIK: A41341: A41341.  
 DR HSSP: P00782: 2SRT.  
 DR MEROPS: S08: UPA;  
 DR Subtilist: BG10591: VPR.  
 DR InterPro: IPR003137: PA.  
 DR InterPro: IPR000209: Peptidase\_S8.  
 DR Pfam: PF00082: Peptidase\_S8; 2.  
 DR Pfam: PF02225: PA; 1.  
 DR PRINTS: PR00723: SUBTILISIN.  
 DR PROSITE: PS00136: SUBTILISIN\_ASP; 1.  
 DR PROSITE: PS00137: SUBTILISIN\_HIS; 1.  
 DR PROSITE: PS00138: SUBTILISIN\_SER; 1.

KW Hydrolyase: Serine protease; zymogen; signal; complete proteome.  
 FT SIGNAL 1 28  
 FT PROPEP 161 160  
 FT CHAIN 161 806  
 FT ACT\_SITE 189 189  
 FT ACT\_SITE 233 233  
 FT ACT\_SITE 534 534  
 SO SEQUENCE 806 AA; 85608 MW; F984E3BF0869DD CRC64;  
 Query Match 9.2%; Score 286; DB 1; Length 806;  
 Best Local Similarity 22.5%; Pred. No. 1.9e-13;  
 Matches 132; Conservative 60; Mismatches 188; Indels 206; Caps 19;  
 OY 156 TRGASLYVKAALDTRKXNKEVO--LKG-----EKIAQXXSNVXYITAKPREKVMN 207  
 DB 87 TKAKNKAIRK--AVKNKVNREVEQVSGFSKMLPANEIPILLAVDVAVVAVPNTYTKDN 144  
 OY 208 DVARGVYKADVAQ-----SYGIGGOGIVAVAPDGLDGR----- 243  
 DB 145 MKDKVYISDVAISFQMDSDAPYIGANDAMDGLTGGIKVAILIDGVEYHNPDLKKNFG 204  
 OY 244 -----NDSSMEAEFRKITALYALGRTNNANDTNGHGHVAGSVLGNCTKNGAP 294  
 DB 205 QYKGYDFVNDYDPEKTPG-----DPRGEATDHGHVAGVAAAG--TIKGVAP 252  
 OY 295 QANLVFQSIIMDSXGGLGLPSNLQTLFSAQASAGARLHTNSMGAANVAYTTDSRNVDY 354  
 DB 253 DATLLAVRVLYPGG--SGTTEVVIAGVRAVODADVMNLISGNSLNPDVATSTAL-DW 309  
 OY 355 VRKNDMTLEFAENGEXPNNGCTISAPGTAKNATVGA-----NLRFPSGSY----- 401  
 DB 310 AMSGVAVATVSNQSGPNQMWVFGSPGTSREASIVGATQPLNEAVTFSSSAKVMGYN 369  
 OY 402 -----ADNT----- 405  
 DB 370 KEDVYALNKNKEVELVEAGIGEAQDFGKDLTGKVAAYVKGSIAPVDKNAKKAGATCM 429  
 OY 406 -----NHAQV 411  
 DB 430 VVYNLNLSELEAVNPMSPPTIKLSLEDEKLYSALKAGETTKTKLYSKRAGEQVADF 489  
 OY 412 SSRGPTKD-GRKPDVMAFGYIIISANSLAPDSSFWANHDSKYAVMGTSWAPPIYVAGN 470  
 DB 490 SSRGPDVMTWIKPDISAGVIVSTIPTDPP-----HPYGGSGQSTSMASPHLAGA 543  
 OY 471 VAOLREHFVKNRGITPKPSL--LKAALAGAA-----DXGLGYPNGNGGVRVTLDRSLNV 524  
 DB 544 VAVIKQ-----AKPKMSVEQIKAMVNTAVTLKSDGGEVYPHNNOGAGSARI---MNA 593  
 OY 525 AYVNESXLSSTOKATYXTATAGPKPLKISLWSDAPASTASVTL 570  
 DB 594 --IKADSLVSPGSYSYGFLEKNGENETFTENQSSIRKSYTL 637  
 RESULT 5  
 SUBT\_BACS9 STANDARD: PRT; 420 AA.  
 AC P28842;  
 DT 01-DEC-1992 (Rel. 24, created)  
 DT 01-DEC-1992 (Rel. 24, last sequence update)  
 DT 15-JUN-2002 (Rel. 41, last annotation update)  
 DE Subtilisin precursor (EC 3.4.21.62).  
 GN SUBI.  
 OS Bacillus sp. (strain TA39).  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OC NCBI\_TaxID=29336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92256481; PubMed-1581352;  
 RA Narinx E., Davail S., Feller G., Gerday C.;  
 RT "Nucleotide and derived amino acid sequence of the subtilisin from  
 Biochim. Biophys. Acta 1131:111-113(1992)."

- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE, IT CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.  
 - CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity for peptide bonds, and a preference for a large uncharged residue in p1. Hydrolyzes peptide amides.  
 - SUBCELLULAR LOCATION: Secreted.  
 - MISCELLANEOUS: STILL ACTIVE AT TEMPERATURES CLOSE TO 0 DEGREES CELSIUS, IT HAS A MARKED HEAT LABILITY.  
 - MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF SPOULATION, AND MANY MUTATIONS WHICH BLOCK SPOULATION AT EARLY STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN IS NOT NECESSARY FOR NORMAL SPOULATION.  
 - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: X62369; CAA44227.1; -  
 PIR: S23407; S23407.  
 DR HSSP: Q93405; IMP.  
 DR MEROPS: S08.0PA; -  
 DR InterPro: IPR000209; Peptidase\_S8.  
 DR Pfam: PF00082; Peptidase\_S8; 1.  
 DR PRINTS: PR00723; SUBTILISIN.  
 DR PROSITE: PS00136; SUBTILISIN\_ASP; 1.  
 DR PROSITE: PS00137; SUBTILISIN\_HIS; 1.  
 DR PROSITE: PS00138; SUBTILISIN\_SER; 1.  
 KW Hydrolyase; Sporulation; Serine protease; Zymogen; Signal.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT PROPEP 27 111  
 FT CHAIN 112 420 SUBTILISIN.  
 FT ACT\_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 182 182 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 360 360 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 SQ SEQUENCE 420 AA; 44086 MW; AEAFL12BD32B26EC CRC64;

Query Match 8.8%; Score 274.5; DB 1; Length 420;  
 Best Local Similarity 28.3%; Pred. No. 5,5e-13;  
 Matches 107; Conservative 56; Mismatches 142; Indels 73; Gaps 21;

QY 168 LDTKQXN--KEYOLRGIEIXIAQXXSNDVXYITAKPE--YKYM---NDVARGIV---KA 216  
 DB 73 MNEKQFNLKRNKLNLTVEVPEL---ETATIDKPEALYNNMAASQSTPMKIKALYNNNS 128  
 QY 217 DVAQSSYGLYGQGVAVADTGLDGRNDSMEHAFRGKITAL--YALGR---NNANDT 271  
 DB 129 SITQTS---GGGGINIAVLDLTGVNTN---HPLLRNNVEOCKDFVTGTYTNNSCFDR 179  
 QY 272 NGHGHVAVSVLNGXGXTK---GMAPOANLVFOSIM--DSXGGLGLSPENLOTLSQXAS 326  
 DB 180 QGHGHVAVSALADGCTGNGYGVAPDLMAYKVLGDGSGSYADDIAAIAIRHAGQDATA 239  
 QY 327 AGARLHTN--SWGAAVNGAYTTDSRVVDVYRKNDMTILFAAGNEXPCNGSTISAPGTAKNA 385  
 DB 240 LNTKYVINMSLSSGSESSILITNAV---YSTNKGVLIIAANSGYQSGISITGPALVANA 296  
 QY 386 ITVGATENIRPSFGSYADNINHVAFSSRGP--KDG---RIKPDVMAFGTYILSARSS 439  
 DB 297 VAVALEN-KVENGTY---RVADFSRSGYSWTDGDVYIQGDVEISAPGAIAYST--- 347  
 QY 440 LAPDSFPAHNSKAYMAGTSMATPIYAGNVAQLREHNVKRGKITPKSL-----L 491  
 DB 348 -----W--FDGGYATISGTSMAASPHAAGLAKIWAQYPSASNDVREGLQYRAYENDI 398  
 QY 492 KALLIAGAAD---XGIGY 506  
 DB 399 LSGYYAGYGDDEFAAGFGF 416

RESULT 6  
 PROA\_XANCP STANDARD; PRT; 580 AA.  
 ID PROA\_XANCP  
 AC P2314;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Extracellular protease precursor (EC 3.4.21.-).  
 GN XCC0851.  
 OS Xanthomonas campestris (pv. campestris).  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xanthomonas.  
 OX NCBI\_TaxID=340;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90251253; PubMed=2187155;  
 RA Liu Y.-N., Tang J.-L., Clarke B.R., Dow J.M., Daniels M.J.;  
 RT "A multipurpose broad host range cloning vector and its use to  
 RT characterise an extracellular protease gene of Xanthomonas campestris  
 RT pathovar campestris.";  
 RL Mol. Gen. Genet. 220:433-440(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33913 / NCPPB 528;  
 RX MEDLINE=22022145; PubMed=12024217;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Canaletto G., Canavan F., Cardoso J., Chambergo F., Clapina L.P.,  
 RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kiraajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.";  
 RL Nature 417:459-463(2002).  
 CC - SUBCELLULAR LOCATION: Secreted.  
 CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: X51635; CAA35962.1; -  
 EMBL: AE012184; AAM40166.1; -  
 DR PIR: S11890; S11890.  
 DR HSSP: P00782; 2SRT.  
 DR MEROPS: S08.0PA; -  
 DR InterPro: IPR000209; Peptidase\_S8.  
 DR Pfam: PF00082; Peptidase\_S8; 1.  
 DR PRINTS: PR00723; SUBTILISIN.  
 DR PROSITE: PS00136; SUBTILISIN\_ASP; 1.  
 DR PROSITE: PS00137; SUBTILISIN\_HIS; 1.  
 DR PROSITE: PS00138; SUBTILISIN\_SER; 1.  
 KW Hydrolyase; Serine protease; Zymogen; Signal.  
 FT SIGNAL 1 32 POTENTIAL.  
 FT PROPEP 33 2136  
 FT CHAIN 2137 580  
 FT ACT\_SITE 177 177 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 237 237 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 409 409 CHARGE RELAY SYSTEM (BY SIMILARITY).



```

FT DISULFID 225 273 BY SIMILARITY.
FT DISULFID 315 352 BY SIMILARITY.
FT DISULFID 450 454 POTENTIAL.
SQ SEQUENCE 580 AA: 57228 MW: 8C9A2CAE4E7F47CB CRC64:

Query Match
Best Local Similarity 25.0%; Score 270; DB 1; Length 580;
Matches 135; Conservative 66; Mismatches 194; Indels 146; Gaps 26;

OY 162 LVKA-XALDFKQNKKEVOLGLEXIXOAXXSDVYIYIAPK-EYKXNDVAPGIVKA--D 217
    ||| ||| : : : : :
DB 99 LVRAERADLRREAEETLMRLQADAPNVQSYEVYQIHLATLPNDRLRSEQAFETNAGLN 158
    ||| ||| : : : : :
OY 218 VAQSSYGLGGGGOIVAVDTG---DGRNDSSMHEFRKXTALXAGRTNAND--- 270
    : : : : : ||| ||| : : : : :
DB 159 IRPAMDATKATGCGIVAVIDGTIISHADLNANITLADYDISATTAARNGRCSNADGCD 218
    : : : : : ||| ||| : : : : :
OY 271 -----TNGRTHVAGSY--LGNGXIN-KGAPQANLYFGSINDXG--- 309
    : : : : : ||| ||| : : : : :
DB 219 WYANEGCAGIPPAASSMWHGTHVAGVAAVNTTGVAGTAYGAKVVPVRLKCGGSLIS 278
    : : : : : ||| ||| : : : : :
OY 310 -----LGLIPSNLQ--TLFSQASGARIHFNMSGAANVAVGAYTTDSRVNDY 354
    : : : : : ||| ||| : : : : :
DB 279 DIADIYVWASGIVSGIPANANPAEYINMSLGGGSGCTTMON-AINGAVSRGT----- 331
    : : : : : ||| ||| : : : : :
OY 355 VRKNDMTILFAAGNEXPNGGTISAPOTAKMAITVGATEN--LRPFGSYADNINHYAOPS 412
    : : : : : ||| ||| : : : : :
DB 332 -----TVYVAAGNDASNVSG-SLPANCANVIAVAATTSAGAKASYENFCTGI----- 377
    : : : : : ||| ||| : : : : :
OY 413 SRGPTKDRIRPDVMAFGRTIISARS--LAPDSEFNHDSKXYAVGGSMTPTIVAGN 470
    ||| ||| : : : : : ||| ||| : : : : :
DB 378 -----DVSAPGSSILSTLNSGTTTPGS-----ASYASVNGTSMASPRIVAG 418
    : : : : : ||| ||| : : : : :
OY 471 VAOLREHVKNKRGITPK--PSLLK--AALIGADMXGIGY-----PNGNQ 511
    ||| : : : : : ||| ||| : : : : :
DB 419 VALVQS--VAPTALITPAAYETLLKNTARALPACGCGGAGIYVAAADVAITAINSGSGG 476
    : : : : : ||| ||| : : : : :
OY 512 CWRGRTLDKSLINVAYNVSSXISTSOAKATYKTTATAGPKPLKLSWSDAPASTTASVTL- 570
    : : : : : ||| ||| : : : : :
DB 477 GGGGNTLTNGTPVTGDC---AATGAELNYITTPAG-----SGTLVPTTS 518
    : : : : : ||| ||| : : : : :
OY 571 --VNDLDLVI--TAPNGTIVGNDFEXKXXNMMDGRNNVNFINKPQSGTITTEVQAY 625
    ||| : : : : : ||| ||| : : : : :
DB 519 GGGGADLYVWASGAPTDASVYTCRPYRS-----GNAETCTITAP-SGTYVYRLKAY 568
    : : : : : ||| ||| : : : : :
OY 626 N 626
DB 569 S 569

RESULT 7
WPR_A_BACSU
ID WPR_A_BACSU STANDARD: PRT: 894 AA.
AC P54423.006726;
DT 01-OCT-1998 (Rel. 34, Created)
DT 13-DEC-1998 (Rel. 37, Last sequence update)
DT 13-JUN-2002 (Rel. 41, Last annotation update)
DE Cell wall-associated protease precursor (EC 3.4.21.-) [Contains: Cell
DE wall-associated polypeptides CWBP23 and CWBP52].
GN WPR_A.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.
RC STRAIN=168;
RX MEDLINE=97158234; PubMed=9004506;
RA Margot P., Karamata D.;
RT "The wprA gene of Bacillus subtilis 168, expressed during exponential
RT growth, encodes a cell-wall-associated protease.";
RL Microbiology 142:3437-3444(1996).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=98015415; PubMed=9353931;
RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
RT "Sequencing of regions downstream of addA (98 degrees) and citC (289
RT degrees) in Bacillus subtilis.";
RL Microbiology 143:3305-3308(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Allion G.,
RA Azevedo V., Bertoro M.G., Bessieres P., Bolotin A., Borchert S.,
RA Bortiss R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Enlian K.D., Errington J., Fabre C., Ferrati E., Foulger D.,
RA Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Hentut A.,
RA Hilbert H., Holtsappel S., Hosono S., Hulio M.F., Ilaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kleerr-Bianhard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogilwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Setiaguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccout E., Takagi T., Takahashi H., Takemaru K.,
RA Taneuchi M., Tamakoshi A., Tanaka T., Terstra P., Tonpori A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassaretto A.,
RA Viri A., Wandelt R., Wedler E., Wedler H., Weitzmeyer K., Yata K.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasuno K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -I- FUNCTION: NOT YET KNOWN. COULD BE INVOLVED IN PROTEOLYCAN
CC DEGRADATION, BY CLEAVAGE OF ITS PEPTIDE BRIDGES.
CC -I- SUBCELLULAR LOCATION: CELL-WALL BOUND.
CC -I- PPM: PROCESSED INTO CWBP23 AND CWBP52.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 58.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U58981; AAC25926.1; -
DR EMBL: Y09476; CAAT0641.1; -
DR EMBL: Z99109; CAB12917.1; -
DR HSSP: Q45670; IDBI.
DR MEROPS: S08.004; -.
DR Subtilist; BG11846; WPR_A.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILINASE.ASP; FALSE_NEG.
DR PROSITE: PS00137; SUBTILINASE.HIS; 1.
DR PROSITE: PS00138; SUBTILINASE.SER; 1.
KW Hydroxylase, serine protease, Cell wall, zymogen, signal;
KW Complete proteome.
FT SIGNAL 1 31
FT CHAIN 32 894 CELL WALL-ASSOCIATED PROTEASE.
FT CHAIN 32 413 POTENTIAL.
FT PROPEP 414 894 CWBP52.
FT CHAIN

```

FT ACT\_SITE 466 466 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 650 650 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CONFLICT 9 9 V -> A (IN REF. 1).  
 FT CONFLICT 14 14 L -> I (IN REF. 1).  
 SQ SEQUENCE 894 AA; 96487 MW; 0F67C33E35F8DBC CRC64;

Query Match 8.4%; Score 262; DB 1; Length 894;  
 Best Local Similarity 23.9%; Pred. No. 1.2e-11;

Matches 134; Conservative 92; Mismatches 211; Indels 124; Gaps 28;

QY 50 FSKOXOTGAALF--LYESENKYLAKGLKLEVPV-----NNKLHIXQFN---G 94  
 DB 245 FDDVSENGASSYKVEKDKAINRLYDKALQSPFLKEIRKQADRLNMKLOQKRTAG 304  
 QY 95 PILEE-----TKOXLEXTGAKITLDYIPDYAYIV---EYEG-----VYXSXXX 134  
 DB 305 ALLTENNIAAKSEVOTT--KVIFKVDNKSLSVHNEMKFSASQSKDISNKKAKKL 362  
 QY 135 IEHVESVEPYLPYXXIDPOLFTKGASLYKAKALDTRKXKEVOLRGIEIXIAQ-XXXXND 193  
 DB 363 FDLNYSFE--LPKDEKONGAYTASAKRVKSAATLTSKMSNVEF---AEFVQEKSLAND 416  
 QY 194 VYXITAKPEKVMNDVARGIVK-ADVA-QSYGLYGOQ-----IVAVADGLDTGRDSS 247  
 DB 417 IQY---PYOMPLKNGEENGKVNADVKYEPANTLTSKRKINDTLIAVDVDTGLAD- 471  
 QY 248 MHEAFRGKITALYA---LGRTNANNDTNGHTVAGSVLD---NGXTNKGMAPOANLVQ 301  
 DB 472 ----LKGKVRTDLGHNFVGNMNNAMDOGHGTIVAGLIAOSONGISMTLNAKAKITPV 527  
 QY 302 SINDSGGLGGLPNSLQTLFSGAXSAGARIHTNSWGAAGVATYDSRVNDVYRKNDMT 361  
 DB 528 KVDLSAG--SGDDEQIALGIKYADKAGAKYINSLG---GGYSRVLEFALKYAADKNVL 581  
 QY 362 ILFAAGNEXNGGTISAPGAKNAITVGATENLRPSRGSYADNINHAQSSSGPTKDR 421  
 DB 582 IAAASGNDGEN--ALSTYPASSKTYMSVGAT-----NKMDDTADFSSNYGKGL--- 625  
 QY 422 IKPDVAPGTXILSARSLAPDSSFWANHDHSKYAYMGTSMAPIYAGVNAQAREHFVK 481  
 DB 626 ---DISAPGSDI---PSLVPNGN-----VTYMGSTMATPYAAAAGLFLAQPKL 670  
 QY 482 RGTTPPSLILKAA--LIAGAAXG-----LGRP-----NGNQGGRVTLDSLN 523  
 DB 671 KRTEVEDMLKKTADDISFESVDGEEELYDDYGDPIEIPKTPGVDMHSGYGRLVNKA 730  
 QY 524 VA----YVNESSXLTSTOKAT 540  
 DB 731 AADLOLKVKNLESTOTAVRGS 751

RESULT 8  
 ELYA\_BACSP  
 ID ELYA\_BACSP STANDARD; PRT; 378 AA.  
 AC P20724;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Alkaline elastase Yab precursor (EC 3.4.21.-).  
 GN ALE  
 OS Bacillus sp. (strain Yab).  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1409;  
 RN [1]  
 RP MEDLINE=89359181; PubMed=2670913;  
 RA Kaneko R., Koyama N., Tsai Y.-C., Juang R.-Y., Yoda K., Yamasaki M.;  
 RT "Molecular cloning of the structural gene for alkaline elastase Yab,  
 a new subtilisin produced by an alkalophilic Bacillus strain.";  
 RL J. Bacteriol. 171:5232-5236(1989).  
 RN [2]  
 RP SEQUENCE OF 111-164.

RA Tsai Y.-C., Lin Y.-T., Li Y.-F., Yamasaki M., Tamura G.;  
 RT "Characterization of an alkaline elastase from alkalophilic Bacillus  
 ya-B.";  
 RL Blochm. Biophys. Acta 883:439-447(1996).  
 CC -1- FUNCTION: DIGEST ELASTIN EFFICIENTLY, HAS A SUBSTRATE PREFERENCE  
 CC FOR ALA IN P1 POSITION.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.  
 CC The European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL, M28537; AA87324.1; -;  
 DR PIR; A33973; A33973.

DR HSSP; Q99405; IMPT.  
 DR MEROPS; S08.00A; -;  
 DR InterPro; IPR000209; Peptidase\_S8.

DR Pfam; PR00082; Peptidase\_S8; 1.  
 DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE\_ASP; 1.  
 DR PROSITE; PS00137; SUBTILASE\_HIS; 1.  
 DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
 KW Hydrolase; Serine protease; Zymogen; Signal.  
 FT SIGNAL 1 27  
 FT PROPEP 28 110  
 FT CHAIN 111 378  
 FT ACT\_SITE 141 141  
 FT ACT\_SITE 171 171  
 FT ACT\_SITE 324 324

FT SEQUENCE 378 AA; 38793 MW; 5A8BD8CC0C62687D CRC64;  
 Query Match 8.3%; Score 257.5; DB 1; Length 378;  
 Best Local Similarity 26.7%; Pred. No. 8.5e-12;  
 Matches 115; Conservative 50; Mismatches 167; Indels 99; Gaps 20;

QY 74 LKKLETVPANNLKHIX-OPENPILBETKXLE--XTGAKITLDYIPDYAYIVEGQVXS 130  
 DB 1 MKRKGIVAGTALITISVAFSSIIAQAEBARKYILGFEQEVMSQ--FVQIDGDEIS 58  
 QY 131 XXXXIEHVESVEPYLPYXXIDPOLFTKGASLYKAKALDTRKXKEVOLRGIEIXIAQXXX 190  
 DB 59 ISSQAEDEVET--DLRHEFDPIVLSVELDPEVDALDLPD-----SN 97  
 QY 191 SNDVYXITAKPEKVMNDVARGI--VKADYAOSSYGLYGOQIVAVADGLDTGRDSSM 248  
 DB 98 --IAYIEEDAETTTQOTYVWGINRVQAPIAOSR-GFTGTVRVAVALDTGI-----SN 146  
 QY 249 HEAFRGKITALYALGRNANNDTNGHTVAGSV--LGNGXTNKGMAPOANLVQSSIM-- 304  
 DB 147 HADLRIRGCASFVPGEP-NISDNGNGHTOVAGTIALINNSIGVLGAPVNDLGYAVVLGA 205  
 QY 305 DSXGGLGGLPNSLQTLFSGAXSAGARIHTNSWGAAGVATYDSRVNDVYRKNDMTIL 364  
 DB 206 SSGSGISGLAQGLQ----MAANNMGMIHANNISLSSAGSATMEQAVN--QATASGLVYVA 258  
 QY 365 AAGNEXPNCGTISAPGTAKNAITVGATE--NLRPSCGSTADNINHAQSSRSGPTDGR 422  
 DB 259 ASGNS--GAGNVGFPARYANMAVAGTDDONNRRATFSOYGAGL----- 299  
 QY 423 KPDVMAPGTXILSARSLAPDSSFWANHDHSKYAYMGTSMAPIYAGVNA----- 472  
 DB 300 --DIVAPGVGV---OSTVPGNG-----YASFNQTSMAPIHAGVAAVLAKQKNPSWS 345  
 QY 473 --OLRHEFYKN 481  
 DB 346 NVQIRNH-LKN 355

```

RESULT 9
AC AQLI_THERMO STANDARD: PRT: 513 AA.
AC P08594;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aqualysin I precursor (EC 3.4.21.-).
GN PSYI.
OS Thermus aquaticus.
OC Bacteria; Thermus/Delnuococcus group; Delnuococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxID=271;
RN [1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 15-23.
RC STRAIN-YT1;
RA MEDLINE=90216674; PubMed=2182621.
RA Teraida I., Kwon S.-T., Miyata Y., Matsuzawa H., Ohta T.;
RT "Unique precursor structure of an extracellular protease, aqualysin
RT I, with NH2- and COOH-terminal pro-sequences and its processing in
RT Escherichia coli."
RL J. Biol. Chem. 265:6576-6581(1990).
RN [2]
RN SEQUENCE OF 75-442 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-YT1;
RA MEDLINE=88225062; PubMed=3286255;
RA Kwon S.-T., Teraida I., Matsuzawa H., Ohta T.;
RT "Nucleotide sequence of the gene for aqualysin I (a thermophilic
RT alkaline serine protease) of Thermus aquaticus YT-1 and
RT characteristics of the deduced primary structure of the enzyme."
RL Eur. J. Biochem. 173:491-497(1988).
RN [3]
RN SEQUENCE OF 128-170.
RA MEDLINE=88151937; PubMed=3162211;
RA Matsuzawa H., Tokugawa K., Hamaki M., Mizoguchi M., Taguchi H.,
RA Teraida I., Kwon S.-T., Ohta T.;
RT "Purification and characterization of aqualysin I (a thermophilic
RT alkaline serine protease) produced by Thermus aquaticus YT-1."
RL Eur. J. Biochem. 171:441-447(1988).
RN [4]
RN FUNCTION: AQUALYSIN I IS A THERMOPHILIC ALKALINE SERINE PROTEASE.
RN THE OPTIMAL TEMPERATURE FOR ITS CASEINOLYTIC ACTIVITY IS 80
RN DEGREES CELSIUS.
RN -1- SUBCELLULAR LOCATION: Secreted.
RN -1- DEVELOPMENTAL STAGE: SECRETED FROM THE EARLY STATIONARY PHASE
RN UNTIL THE TIME THE CELLS CEASE TO GROW.
RN -1- PTM: THE N- AND C-TERMINAL PRO-SEQUENCES ARE REMOVED THROUGH THE
RN PROTEOLYTIC ACTIVITY OF AQUALYSIN I ITSELF, IN THAT ORDER, THE
RN C-TERMINAL PRO-SEQUENCE IS REQUIRED FOR TRANSLLOCATION OF THE
RN PROTEASES ACROSS THE OUTER MEMBRANE.
RN -1- PTM: TWO DISULFIDE BONDS ARE PRESENT.
RN -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 58.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D90108; BAA1435.1; .
DR EMBL, X07734; CAA30559.1; .
DR PIR, S00620; S00620.
DR PIR, A35742; A35742.
DR HSSP, P06873; 2PRK.
DR MEROPS, S08.051; .
DR InterPro, IPR000209; Peptidase-S8.
DR Pfam, PF00082; Peptidase-S8; 1.
DR PRINTS, PRO0723; SUBTILISIN.
DR PROSITE, PS00136; SUBTILASIN.ASP; 1.
DR PROSITE, PS00137; SUBTILASE_HIS; 1.
DR PROSITE, PS00138; SUBTILASE_SRS; 1.
DR Hydrolase; Serine protease; Zymogen; Signal.
KW

```

```

FT SIGNAL 1 14
FT PROPEP 15 127
FT CHAIN 128 408 AQUALYSIN I.
FT PROPEP 409 513
FT ACT_SITE 166 166 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 197 197 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 349 349 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 513 AA; 53913 MW; DDPEED6DA508785 CRC64;

Query Match 8.2%; Score 255; DB 1; Length 513;
Best Local Similarity 24.8%; Pred. No. 1.9e-11;
Matches 133; Conservative 51; Mismatches 161; Indels 192; Gaps 29;

OY 146 PXXII-DPO-----LFTGASLYKAKALDTKXKNEVQ-----LRGI---182
DB 41 PVTGLDDPEALPERITVFKKQSGSLDGGITTLQRLRPAQGVVYTAQYIGALDGFRAE 100
OY 183 ---EXIAQXXXSNDVXYTTAKPERKVMNDVARGIVADVAQS-----221
DB 101 MAPQALEAFROSPDEFLEADK-----VVRAMATQSPAPMGRLDRIDRPLSN 149
OY 222 --SYGLYGQGIYAVADTGLDTGRNDSMHEAFRKITALY-ALGRTNANDTNGCHTV 278
DB 150 SYTYATATGRGVNVVYIDGIRT-----THREFGRRARYGDALG--GNQDDCNHGCHTV 201
OY 279 AGSVLGGXGNTNKGMPAQNLYFOSTMD--SXGGLGGLPSNLTQLFSQAXSAGARLHTN-- 334
DB 202 AGTI---GGVTVGVAKAVNLAVRVLDCNCGSGTSGVIAGVDMV-----TNNHRRPA 250
OY 335 ----SNGCAVNAATYTTDSRNNDDVVRKN---DMTILFAGCNEXPCGTISAPRAKNAT 387
DB 251 VANNSLGGVSTRA-----LDNAVKNSIAAGVYAVAAANDNANACNTS-PARYAALIT 302
OY 388 VGAT--ENLRPFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTXILSARSLAPDS 445
DB 303 VGATSSDAARSPNSNGSCV-----DLRPAASIPSA-----334
OY 446 FMANHDSKYAVMGKTSMAPIYAGNVQALREHVKNGRTGP--RPSLKAALLAGADX 502
DB 335 -WYTSPTAQTQLTNGISMATPHVAG-VAAL--YLEONPSATPASPASAILNGATGRLSGI 390
OY 503 GLGVPNGCGWGVTLDSLVNVAAYNESSXLSOKATYXFTATAGPLKLSLWSDPA 562
DB 391 GSGSPN-----RLVY---SLSSGSGST-----APC 413
OY 563 ST---TASVTLVNDLVLVTAPNGTXY-----VGNDPXXPPXXNMDG 601
DB 414 TCSYITGSLSGRGDYNF---QPNCTIYYSPAGTHRAMLRGPACTDFDL-YLMNRDG 466

RESULT 10
ELIA_BACAO STANDARD: PRT: 380 AA.
ID ELIA_BACAO
AC P27693;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alkaline protease precursor (EC 3.4.21.-).
OS Bacillus alcaicophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1445;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-PB92;
RA MEDLINE=91282483; PubMed=2059048;
RA van der Laan J.C., Gerritse G., Mulleners L.J.M., van der Hoek R.A.,
RA Quax W.J.;
RT "Cloning, characterization, and multiple chromosomal integration of a
RT Bacilli alkaline protease gene."
RT Appl. Environ. Microbiol. 57:901-909(1991).
RN [2]
RN X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RC STRAIN-PB92;

```

```

RX MEDLINE=92390330; PubMed=1518788;
RA van der Laan J.C., Teplyakov A.V., Kelders H., Kalk K.H., Misset O.,
RA Mulleners L.J.M., Dijkstra B.W.;
RT "Crystal structure of the high-alkaline serine protease PB92 from
RT Bacillus alcalophilus.";
RL Protein Eng. 5:405-411(1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RX MEDLINE=93078250; PubMed=1447775;
RA Sobek H., Hecht H.-J., Aehle W., Schomburg D.;
RT "X-ray structure determination and comparison of two crystal forms of
RT a variant (Asn115Arg) of the alkaline protease from Bacillus
RT alcalophilus refined at 1.85-A resolution.";
RL J. Mol. Biol. 228:108-117(1992).
RN [4]
RP STRUCTURE BY NMR OF 112-380.
RX STRAIN=PB92;
RX MEDLINE=97277237; PubMed=9115441;
RA Martin J.R., Mulder F.A., Karimi-Nejad Y., van der Zwan J.,
RA Martini M., Schipper D., Boelens R.;
RT "The solution structure of serine protease PB92 from Bacillus
RT alcalophilus presents a rigid fold with a flexible substrate-binding
RT site.";
RL Structure 5:521-532(1997).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M65086; AAA22212.1; -
DR EMBL: A13738; CA01128.1; -
DR PIR: A49778; A49778.
DR PDB: 1AH2; 15-APR-98.
DR MEROPS: S08_038; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILASIN.
DR PROSITE: PS00136; SUBTILASIN_ASP; 1.
DR PROSITE: PS00137; SUBTILASIN_HIS; 1.
DR PROSITE: PS00138; SUBTILASIN_SER; 1.
KW Hydroxylase; Serine protease; Zymogen; Signal; 3D-structure.
FT SIGNAL 1 27
FT PROPEP 28 112
FT CHAIN 113 380
FT ACT_SITE 143 143 ALKALINE PROTEASE.
FT ACT_SITE 173 173 CHARGE RELAY SYSTEM.
FT ACT_SITE 326 326 CHARGE RELAY SYSTEM.
SQ SEQUENCE 380 AA; 38853 MW; 539EA72771B6682C CRC64;
Query Match 8.1%; Score 250.5; DB 1; Length 380;
Best Local Similarity 24.4%; Pred. No. 2.8e-11;
Matches 123; Conservative 57; Mismatches 148; Indels 177; Gaps 22;

```

```

Db 99 -----ISYIEDAEVTTMAQSVPGWISRWQAPAAHNHGLTGSQVKAVALDTGIST 148
Qy 242 GRNDSSMEHAFRGKITATYALGRTNANDTNGHTHVAGSV--LNGXTNKGMAPOANLY 299
Db 149 -----HPDLNIRGASVPEEP-STQDGNHGHVACTIALNNSIGVGVANAEIX 200
Qy 300 FQSIDSKXGGLGDLPSNLOTLEFSQAXSAGARIHTNSWGA-----AVNGAYTTDSRN 350
Db 201 AKVVLGASG--SGSVSSIKQGLEMGAGNNMGHVAANLSLGSPPSATLEQAVNSA---TSRG 255
Qy 351 VDDYRKNDMTLFPAGNEXRPNGTISAPGTAKNATVGTATE--NLRPSFGSYADINHY 408
Db 256 V-----LVVAASGNS--GAGSISTPARYANMAVAGATQNNRASFSGYAGGL---- 301
Qy 409 AOFSSRGPTKDGRIKPDVWAPGTXILISARSSLAPDSFMANHDSKRYAVGTSMTAPTYA 468
Db 302 -----DIVAPGVNVQSTYPG-----STYASLNGTSMATPHYA 333
Qy 469 GNVV-----QLREHFVN 481
Db 334 GAALVVKRKPMSMVQIRNH-LKN 357
RESULT 11
ELVA_BACCS
ID ELVA_BACCS STANDARD: PRT; 380 AA.
AC P41362;
DT 01-NOV-1995 (Rel. 32; Created)
DT 01-NOV-1995 (Rel. 32; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Alkaline protease precursor (EC 3.4.21.-).
OS Bacillus clausii.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=79980;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
RX MEDLINE=93043753; PubMed=1368952;
RA Takami H., Kobayashi T., Kobayashi M., Yamamoto M., Nakamura S.,
RA Aono R., Horikoshi K.;
RT "Molecular cloning, nucleotide sequence, and expression of the
RT structural gene for alkaline serine protease from alkaliphilic
RT Bacillus sp. 221.";
RL Biosci. Biotechnol. Biochem. 56:1455-1460(1992).
RN [2]
RP SEQUENCE OF 112-129.
RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
RA Horikoshi K.;
RL (In) Horikoshi K. (eds.);
RL Microorganisms in alkaline environments, pp.187-194, VCH,
RL Weinheim (1991).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S48754; AAC60420.1; -
DR EMBL: D13157; BAA02442.1; -
DR EMBL: A26817; CA01836.1; -
DR EMBL: A22550; CA01611.1; -
DR HSSP: P29600; 1GCI.
DR MEROPS: S08_103; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILASIN.
DR PROSITE: PS00136; SUBTILASIN_ASP; 1.
DR PROSITE: PS00137; SUBTILASIN_HIS; 1.

```

DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
 KW Hydrolyase; Serine protease; zymogen; signal.  
 FT SIGNAL 1 27  
 FT PROPEP 28 111  
 FT CHAIN 112 380  
 FT ACT\_SITE 143 143  
 FT ACT\_SITE 173 173  
 FT ACT\_SITE 326 326  
 SO SEQUENCE 380 AA; 38826 MW; 5F73ABC68D56831 CRC64;

Query Match 8.1%; Score 250.5; DB 1; Length 380;  
 Best Local Similarity 24.4%; Pred. No. 2.8e-11;  
 Matches 123; Conservative 57; Mismatches 148; Indels 177; Gaps 22;

OY 5 KKEFLSLVSAALISTYALKNSAGAKRFLDRKGIOTTTDXKGSKQXQGAALFVE 64  
 DB 2 KKLGLKIVASTALISAFSSSTIASAEBAEKY-----LGFNEQ----- 42  
 OY 65 SENVKLKKGLKLETPANNKLIHQFNGLLEETKQIXLTXGAKILIDYIPYAYVEY 124  
 DB 43 -----EAVSEFEQVEANDEVAL-----LSEEEVELE-----LHMF 75  
 OY 125 EG-DVXSXXXXIEHVESPEPYLPYXIDPOLFTKGASXLVKAALDTPKQXKEVQLKIE 183  
 DB 76 ETIVLVSTELSPEDVDALE-----LDPA----- 98  
 OY 184 XIAQXXXSNVXYITAKPEPKV-NDVARGIYKADY-AQSSYGLYGOGIVAVADTGLDT 241  
 DB 99 -----ISYIEDAEVTTMAQSVPMGISRVOAPAHNRGLTSGGVAVVLDIGIST 148  
 OY 242 GRNDSMHEAFRKITALYALGRNNANDTNGHVAAGV--LGNCXTNKGMARPOLV 299  
 DB 149 -----HPLDNRGASFSYRPEP-STODGNCHGTNAGTIALNLSIGLVGAAPSALY 200  
 OY 300 FQSIDMXKGLGLPSNLQTLFSQAXSAGARIHTNSMCA-----AVNGAYTTDSRN 350  
 DB 201 AVKVLGASG--SGSVSSIAQGLEMGNGNMHVALISLSSPEPSATLEQANSL---TSSG 255  
 OY 351 VDDYRKNDMTILFAAGNEXPGNGTISAPGTAKNAITVGATE-NLRPSFGSYADININ 408  
 DB 256 V-----LVVAASGNS--GAGSISYPARVANMAAGATDNNNNASYSQAGNL--- 301  
 OY 409 AOFSSRGPTDKGRIPDVAPGTXTILSARSSLAPDSSFMNHDPSKYAYMGTSMAPTIVA 468  
 DB 302 -----DIYAPGVNVQSTYPG-----STVASLNGTSMATPHVA 333  
 OY 469 GNVA-----QLREHFVK 481  
 DB 334 GAALVYKOKNPSWMSVQIRNH-LKN 357

RESULT 12  
 ALP\_CEPAC STANDARD: PRT: 402 AA.  
 ID SUBN\_BACNA  
 AC P29118;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Alkaline proteinase precursor (EC 3.4.21.-) (ALP).  
 GN ALP.  
 OS Cephalosporium acremonium (Acremonium chrysogenum).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Hypocreales; Hypocreaceae; mitosporic Hypocreaceae; Acremonium.  
 OX NCBI\_TaxID=5044;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91299283; PubMed=1368696;  
 RA Isozaki T., Fukagawa M., Kojima H., Kotsaka M., Aoki H., Imanaka H.;  
 RT "Cloning and nucleotide sequences of the complementary and genomic  
 DNAs for the alkaline protease from Acremonium chrysogenum.";  
 RL Agric. Biol. Chem. 55:471-477(1991)  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D00923; BA00765.1; -.  
 DR PIR; J00332; J00332.  
 DR HSSP; P06873; 2PKK.  
 DR MEROPS; S08.00A; Peptidase-S8.  
 DR InterPro; IPR000209; Peptidase-S8.  
 DR Pfam; PF00082; Peptidase-S8.  
 DR PRINTS; PR00723; SUBTILISIN.  
 DR PROSITE; PS00136; SUBTILASE\_ASP; 1.  
 DR PROSITE; PS00137; SUBTILASE\_HIS; 1.  
 DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
 KW Hydrolyase; Serine protease; zymogen; signal.  
 FT SIGNAL 1 20  
 FT PROPEP 21 120  
 FT CHAIN 121 402  
 FT ACT\_SITE 160 160  
 FT ACT\_SITE 191 191  
 FT ACT\_SITE 347 347  
 SO SEQUENCE 402 AA; 42099 MW; 8D030CCD42D918E1 CRC64;

Query Match 7.9%; Score 246.5; DB 1; Length 402;  
 Best Local Similarity 30.3%; Pred. No. 5.9e-11;  
 Matches 91; Conservative 33; Mismatches 101; Indels 75; Gaps 14;

OY 227 GQGIYAVADTGLDTRNDSMHEAFRK-ITALYALGRNNANDTNGHVAAGVSGN 285  
 DB 151 GSGTYAVADTGL-----LESHNEFSGRALTYNAVGSN--ADNNGHTVAGTI--- 199  
 OY 286 GXTNKGAPOANLVFQSIDMXSG-----LGLPSNLQTLFSQAXSAGARI-----HT 333  
 DB 200 GGRTYGVAKNLILIAVKYFRGSSSTSIILDFMVAANDTIINRGONKAISMSLGGYS 259  
 OY 334 NSMGAAVNGAYTTDSRNVDYRKNDMTILFAAGNEXPGNGTISAPGTAKNAITVG--AT 391  
 DB 260 SAFNNAVNTAVY---SRGVLVYV-----AAGNDNQNNANYS-PASANAITYGSIAS 306  
 OY 392 ENLRPSFGSYADININVAOFSSRGPTDKGRIPDVAPGTXTILSARSSLAPDSSFWAHD 451  
 DB 307 NMARSSFSNYSVL-----DIFAPGTSILSA-----WIGGN 337  
 OY 452 SKYAYMGTSMAPTIVAGNVAVQLREHFVKNRGTPKPSILKALIA---GAADXLGYPN 508  
 DB 338 SATNTISGTSMAPTHTGVVLYLQ-----ALEGLTSGAARLALATITGRVSNPSSGSPN 393

RESULT 13  
 SUBN\_BACNA STANDARD: PRT: 381 AA.  
 ID SUBN\_BACNA  
 AC P35835;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Subtilisin NAT precursor (EC 3.4.21.62).  
 GN APRN.  
 OS Bacillus subtilis var. nato.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=86029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NC2-1;  
 RX MEDLINE=93113095; PubMed=1369081;  
 RA Nakamura T., Yamagata Y., Ichishima E.;  
 RT "Nucleotide sequence of the subtilisin NAT gene, aprN, of Bacillus  
 RL biostr. Biotechnol. Biochem. 56:1869-1871(1992).  
 CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE.





Db 287 -----VDSNQKASFSVGP-----ELDYMAPGVSIQSTLPG-----NKYGA 323  
Oy 457 MGMTSMATPIVAGNYA 472  
Db 324 YNGTSMASPHVAGANA 339

RESULT 15  
SUBT\_BACLI STANDARD: PRT: 379 AA.  
AC P00780.  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Subtilisin Carlsberg precursor (EC 3.4.21.62).  
OS APR.  
OS Bacillus licheniformis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1402;  
RN [1]  
RP SEQUENCE FROM N.A..  
RC STRAIN=NCIB 6816.  
RA MEDLINE=86093688; PubMed=3001653;  
RA Jacobs M., Eliasson M., Uhlen M., Flock J.-I.;  
RT "Cloning, sequencing and expression of subtilisin Carlsberg from  
RT Bacillus licheniformis."  
RL Nucleic Acids Res. 13:8913-8926(1985).  
RN [2]  
RP SEQUENCE OF 106-379.  
RX MEDLINE=68234702; PubMed=4967581;  
RA Smith E.L., Delange R.J., Evans W.H., Landon M., Markland F.S.;  
RT "Subtilisin Carlsberg. V. The complete sequence; comparison with  
RT subtilisin BPN'; evolutionary relationships."  
RL J. Biol. Chem. 243:2184-2191(1968).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF VARIANT WITH SELENOCYST-325.  
RX MEDLINE=93291170; PubMed=8512925;  
RA Syed R., Wu Z.P., Hogle J.M., Hilvert D.;  
RT "Crystal structure of selenosubtilisin at 2.0-A resolution."  
RL Biochemistry 32:6157-6164(1993).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 106-379.  
RX MEDLINE=98087517; PubMed=9425066;  
RA Stoll V.S., Eger B.T., Hynes R.C., Martichonok V., Jones J.B.,  
RA Pai E.F.;  
RT "Differences in binding modes of enantiomers of 1-acetamido boronic  
RT acid based protease inhibitors: crystal structures of gamma-  
RT chymotrypsin and subtilisin Carlsberg complexes."  
RL Biochemistry 37:451-462(1998).  
CC -I- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,  
CC IF CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.  
CC -I- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity  
CC for peptide bonds, and a preference for a large uncharged residue  
CC in pI. Hydrolyzes peptide amides.  
CC -I- SUBCELLULAR LOCATION: Secreted.  
CC -I- BIOTECHNOLOGY: Used as a detergent protease. Sold under the name  
CC Alcalase by Novozymes.  
CC -I- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF  
CC SPOROULATION. AND MANY MUTATIONS WHICH BLOCK SPOROULATION AT EARLY  
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN  
CC IS NOT NECESSARY FOR NORMAL SPOROULATION.  
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X03341; CAB56500.1; -.

DR PIR; A00968; SUBSC.  
DR PIR; A24111; SUBSCL.  
DR PDB; 1CSE; 15-OCT-89.  
DR PDB; 1SCA; 31-JAN-94.  
DR PDB; 1SCB; 31-JAN-94.  
DR PDB; 1SCD; 31-JAN-94.  
DR PDB; 1SEC; 15-JAN-95.  
DR PDB; 2SEC; 15-JAN-95.  
DR PDB; 1SCN; 31-AUG-94.  
DR PDB; 1AF4; 16-JUN-97.  
DR PDB; 1SEL; 31-OCT-93.  
DR PDB; 1YSB; 18-MAR-98.  
DR PDB; 3VSB; 25-MAR-98.  
DR PDB; 1AVT; 25-MAR-98.  
DR PDB; 1AV7; 01-APR-98.  
DR PDB; 1BE6; 14-OCT-98.  
DR PDB; 1BE8; 13-JAN-99.  
DR PDB; 1BEK; 18-NOV-98.  
DR PDB; 1BFU; 18-NOV-98.  
DR MEROPS; S08.001; -.  
DR InterPro; IPR000209; Peptidase\_S8.  
DR Pfam; PF00082; Peptidase\_S8; 1.  
DR PRINTS; PR00723; SUBTILISIN.  
DR PROSITE; PS00136; SUBTILASE\_ASP; 1.  
DR PROSITE; PS00137; SUBTILASE\_HIS; 1.  
DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
KW Hydrolyase; Sporulation; Serine protease; Zymogen; Signal;  
KW 3D-structure.  
FT SIGNAL 1 29  
FT PROPEP 30 105  
FT CHAIN 106 379  
FT ACT\_SITE 137 379  
FT ACT\_SITE 168 168  
FT ACT\_SITE 325 325  
FT ACT\_SITE 207 207  
FT CONFLICT 233 233  
FT CONFLICT 262 265  
FT CONFLICT 316 316  
FT TURN 111 111  
FT TURN 112 115  
FT HELIX 116 117  
FT TURN 118 122  
FT HELIX 123 125  
FT TURN 129 130  
FT TURN 132 137  
FT STRAND 142 143  
FT TURN 145 146  
FT STRAND 149 154  
FT TURN 157 158  
FT TURN 161 162  
FT TURN 168 177  
FT HELIX 190 191  
FT TURN 193 198  
FT STRAND 202 203  
FT TURN 208 220  
FT HELIX 221 222  
FT TURN 222 222  
FT STRAND 225 228  
FT STRAND 232 232  
FT HELIX 237 248  
FT TURN 249 250  
FT STRAND 252 256  
FT STRAND 263 263  
FT TURN 264 265  
FT STRAND 266 266  
FT TURN 271 271  
FT STRAND 272 274  
FT TURN 276 277  
FT STRAND 279 284  
FT TURN 286 287  
FT STRAND 290 290  
FT TURN 292 293  
FT TURN 298 299  
FT STRAND 302 305



```
FT STRAND 309 313
FT TURN 314 316
FT STRAND 317 321
FT HELIX 324 341
FT TURN 343 344
FT HELIX 347 356
FT STRAND 359 359
FT HELIX 364 367
FT TURN 368 369
FT STRAND 371 371
FT HELIX 374 377
FT TURN 378 378
SQ SEQUENCE 379 AA; 38908 MW; F19A6DC5761FB504 CRC64;
```

```
Query Match 7.8%; Score 242.5; DB 1; Length 379;
Best Local Similarity 23.4%; Pred. No. 1.1e-10;
Matches 118; Conservative 55; Mismatches 158; Indels 173; Gaps 19;

QY 1 MRKKKVFSLVSAAILSTVALXNPSAGX--ARKFDLDFKGIQTITDXXGFSKQXGTGA 58
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MMRKKSFMGLMTAFMLVFTMAFSDSASAOAPAKNVEKDY-----IVGFSGVKTAS 52

QY 59 AAFVSEBNVKKLKKLKKLKVTPANNKLIHXFNGPILEETKQXLEXTGAKILDIYIPDY 118
   | : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 53 ----VKKDIIKESGKVDK-----QFR--IINAKAKLDKEALKEVKNNDPDV 93

QY 119 AYIVEYEGDVXSXXXXXIEHVESVEPYLPXYXIDPOLFTKGASXLYKAXALDTKQANKEVQ 178
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 94 AYVEE-----DIVAHALAQVTPYGIIP----- 114

QY 179 LRGIEXIAQXXXNDXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGOGQIVAVADTG 238
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 115 -----LIRADKVQAO-GFKGANVKVAVALDTG 139

QY 239 LDTGRNDSMHEAFRKITALYALGRTNNANDTNGHGTIVAGSV--LGNGXTNKGMAPQA 296
   | : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 140 IQASHPDLNVVGG-----ASFVAGEAYN-TDGNHGHTHVAGTVAALDNTTGLVGAPSV 192

QY 297 NLVFQSIMDSXGGLGIPSNLOTLFSSQAXSAGARIHTNSMGA-----AVNGAYTTD 347
   | : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 193 SLVAVKVLNMSG--SGTYSGIIVSGIEMATNTGMDVINNLSLGPSGSTRAMQAVDNAYA-- 248

QY 348 SRNVDDYVRKNDMTILFAAGNEXPNG--GTISAPGTAKNAITVGATENLRPSFGSYADNI 405
   | : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 249 -----RGVVVVAAGNSGSGNTTIGYPAKYDSVIAYGA-----VDSN 287

QY 406 NHVAFSSRGPTDGRKIKPDVAPGTXILSARSSLAPDSSFWANHDSKAYVMGTSMATP 465
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 288 SNRASFSSTVG-----AELVMAFGAGVSTYPT-----STYATLNGTSMASP 329

QY 466 IVAGNVA-----QLREHFVKNR 482
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 330 HVAGAALILSKHPNLSASQVRNR 353
```

Search completed: April 1, 2003, 12:38:52  
Job time : 15.0078 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 12:35:29; Search time 60.5473 Seconds  
(without alignments)  
2177.965 Million cell updates/sec

Title: US-09-920-954-2

Sequence: 1 MRKKKRVLSVSAAILST.....EVOANVPVGPQESLAIVN 640

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SP\_ARCHAEA:\*
- 2: SP\_BACTERIA:\*
- 3: SP\_FUNGI:\*
- 4: SP\_HUMAN:\*
- 5: SP\_INVERTEBRATE:\*
- 6: SP\_MAMMAL:\*
- 7: SP\_MHC:\*
- 8: SP\_ORGANELLE:\*
- 9: SP\_PHAGE:\*
- 10: SP\_PLANT:\*
- 11: SP\_RODENT:\*
- 12: SP\_VIRUS:\*
- 13: SP\_VERTEBRATE:\*
- 14: SP\_UNCLASSIFIED:\*
- 15: SP\_VIRUS:\*
- 16: SP\_BACTERIAP:\*
- 17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	2155	69.5	434	2	Q9AQR3
2	2155	69.5	434	2	Q9AQR3
3	2082	67.1	434	2	Q9AQR0
4	1952.5	62.9	433	2	Q9AQR1
5	1948.5	62.8	433	2	Q9AQR4
6	1941.5	62.6	433	2	Q9AQR2
7	493	15.9	1825	5	Q879W1
8	478	15.4	1702	5	Q9GTN7
9	412	13.3	561	16	Q8RBJ2
10	408	13.2	654	17	Q8UOC9
11	391	12.6	1239	16	Q9FBZ4
12	368.5	11.9	1253	16	Q9FC06
13	356	11.5	1102	2	P95684
14	350.5	11.3	444	16	Q9KJ57
15	337	10.9	1245	16	Q9RL54
16	336.5	10.8	412	2	Q9AER6

17	336.5	10.8	412	16	Q8RC68	Q8RC68 thermomaneer
18	321.5	10.4	1220	16	Q9LOA0	Q91060 streptomyc
19	319.5	10.3	442	16	Q31788	Q31788 bacillus su
20	315	10.2	824	2	Q45464	Q45464 thermococu
21	310	10.0	891	1	Q93635	Q93635 thermococu
22	301.5	9.7	1398	1	Q9P9L1	Q9P9L1 pyrococcus
23	295.5	9.5	431	2	Q9S3P6	Q9S3P6 bacillus sp
24	281.5	9.1	434	2	Q54327	Q54327 bacillus sp
25	280	9.0	1135	1	Q9P9P1	Q9P9P1 uncultured
26	279.5	9.0	419	2	Q45681	Q45681 bacillus su
27	276.5	8.9	799	16	Q9KEM1	Q9KEM1 bacillus su
28	272	8.8	397	2	P97097	P97097 bacillus sp
29	270	8.7	621	2	Q9F486	Q9F486 alteromonas
30	270	8.7	621	2	Q53401	Q53401 alteromonas
31	269	8.7	379	2	Q66153	Q66153 bacillus sp
32	268.5	8.7	715	2	P70765	P70765 alteromonas
33	260.5	8.4	1345	1	Q54437	Q54437 staphylothe
34	257.5	8.3	403	2	Q45463	Q45463 bacillus sp
35	257	8.3	629	2	Q93RG8	Q93RG8 alteromonas
36	255	8.2	692	2	Q9EXK0	Q9EXK0 bacillus ps
37	255	8.2	1098	16	Q9L128	Q9L128 streptomyc
38	250.5	8.1	378	2	Q45466	Q45466 bacillus sp
39	250.5	8.1	601	2	Q45540	Q45540 bacteroides
40	250	8.1	382	2	Q45522	Q45522 bacillus sp
41	249.5	8.0	757	16	Q9K6G6	Q9K6G6 bacillus ha
42	248	8.0	627	16	Q9RUD0	Q9RUD0 delnecoccus
43	246.5	7.9	467	2	Q93104	Q93104 xanthomonas
44	246	7.9	537	16	Q9K3X9	Q9K3X9 streptomyc
45	243.5	7.8	374	2	Q9FP43	Q9FP43 bacillus 11

ALIGNMENTS

RESULT 1

ID Q9AQR3 PRELIMINARY: PRT: 434 AA.

AC Q9AQR3; 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE Protease (Fragment).

GN PROA.

OS Bacillus sp. 9860.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

OX NCBI\_TaxID=133778;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=9860.

RX MEDLINE=20568675; PubMed=11118284;

RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,

RA Horikoshi K.;

RT "Novel oxidatively stable subtilisin-like serine proteases from

RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and

RT evolutionary relationships.";

RL Biochem. Biophys. Res. Commun. 279:313-319(2000).

DR EMBL; AB046403; BAB21266.1; .

DR HSSP; P00782; 1SUP.

DR InterPro; IPR000209; Peptidase\_S8.

DR Pfam; PF00082; Peptidase\_S8; 2.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00137; SUBTILASE\_HIS; UNKNOWN\_1.

DR PROSITE; PS00138; SUBTILASE\_SER; 1.

FT NON\_TER

FT TER

SO SEQUENCE 434 AA; 45311 MW; AFC9F78BBI43527E CRC64;

Query Match 69.5%; Score 2155; DB 2; Length 434;

Best Local Similarity 96.3%; Pred. No. 2; Ie-122;

Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 207 NDVARGIVKADVAQSSYGLGGQIVAAVDGTGGRNDSMHEARFGKITLALGRTN 266

D	b	1	NDVARGITKADVAASSVGLVGQGIIVAVADTGIDTGRNDSMHFAFGKITALLYALGRTN	60
Q	y	267	NANDNGHGTTHVAGSVLNGXGXTNKGMAPQANLYFQSIIMDSXGLGLPSNLQTLFESQAXS	328
D	b	61	NANDNGHGTTHVAGSVLNGXGXTNKGMAPQANLYFQSIIMDSXGLGLPSNLQTLFESQAFS	120
Q	y	327	AGARITHNSKCAAVNGAYTTDSRVDVYKRNMDTILFPAAGNEKPNCGTISAPETAKNAI	386
D	b	121	AGARITHNSKCAAVNGAYTTDSRVDVYKRNMDTILFPAAGNEKPNCGTISAPETAKNAI	180
Q	y	387	TVGATENTLRPFEGSGYADININVAOFSSRGPKDKDRIRKPDVMAPECTXILISARSSLAPRSSF	446
D	b	181	TVGATENTLRPFEGSGYADININVAOFSSRGPKDKDRIRKPDVMAPECTXILISARSSLAPRSSF	240
Q	y	447	WANHDSKAYVAGGYSMAPIVAGVNAOLREHFVYVNRGITPKSPILSKAALIAAGADKSLGY	506
D	b	241	WANHDSKAYVAGGYSMAPIVAGVNAOLREHFVYVNRGITPKSPILSKAALIAAGADVGLGY	300
Q	y	507	PNGNOGWGRVTLDDKSLNVAIVNESSXLISTSOKATYFTATAGPKDLKISLWSDAPASTTA	566
D	b	301	PNGNOGWGRVTLDDKSLNVAIVNESSALSTSOKATYFTATAGPKDLKISLWSDAPASTTA	360
Q	y	567	SVTLVNDLDLVITLRPNGTXYVGNDFXXAPXXXNMWDGRNNVEVFLINXQSGITYTTEVOAYN	626
D	b	361	SVTLVNDLDLVITLRPNGTXYVGNDFSAFDMNMDGRNNVEVFLINXQSGITYTTEVOAYN	420
Q	y	627	VPVGPQXFSLAIYN 640	
D	b	421	VPVGPQXFSLAIYN 434	

## RESULT 2

AD	0930V9:	PRELIMINARY;	PRT;	434	AA.
AC	0930V9:				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	Protease (Fragment).				
GN	PROF.				
OS	Bacillus sp. KSM-KP43.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales				
OC	Bacillaceae; Bacillus.				
OX	NCBI_TaxID=109322;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=KP43;				
RA	Itch S., Saeki K.;				
RT	"new protease.";				
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AB051423; BAB5674.1; -. S8.				
DR	InterPro; IPR000209; Peptidase_S8.				
DR	Pfam; PF000082; Peptidase_S8; 1.				
DR	PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.				
DR	PROSITE; PS00138; SUBTILASE_SER; UNKNOWN_1.				
FT	NON_TER	1			
FT	NON_TER	1			
FT	NON_TER	434			
FT	NON_TER	434			
SQ	SEQUENCE	434	AA:	45302 MW;	0B08418191853GC3 CRC64;

Query Match	69.58	Score 2155	DB 2	Length 434
Best Local Similarity	96.38	Pred. No. 2.1e-122		
Matches 418; Conservative	0	Mismatches 16	Indels 0	Gaps 0

QY	207	NDVARGIKADVAOSSSYGLVGGQGIIVADVADTGTGRNDSMHEAFGKITALYALRTN	266
Db	1	NDVARGIKADVAOSSSYGLVGGQGIIVADVADTGTGTGRNDSMHEAFGKITALYALRTN	60
QY	267	NANNTNGHGTVAASVYLGNGKXTNKGMAPOANLYVQSIIMDSXGIGIGIPSNLOITLFFSQAXS	326
Db	61	NANNTNGHGTVAASVYLGNGSTNKGMAPOANLYVQSIIMDSGGIGIGIPSNLOITLFFSQAYS	1220
QY	327	AGARIHTNSMCAAVNAYTTDSRNVDYVYRKNNDYTLFPAAGNEXPNGGTSISAPGTARKAI	386

Db	121	AGARIHTSMGWAANVGAATTDTSRNVDDYVRKNDMTILFEAAGNEBPNGGTISAPCTAKNAI	180
QY	387	TVGATENTLRPESEGYADNININVAOFSSRGPCKDGRIRFDYMAPOCTXILSRSSLAPDSSF	446
Db	181	TVGATENTLRPESEGYADNININVAOFSSRGPCKDGRIRFDYMAPOCTXILSRSSLAPDSSF	240
QY	447	WANHSKTAAYWGGTSMATPIYAGVNAOOLREHEFVNKRGITPRPSLLKALLAGAADKGLG	506
Db	241	WANHSKTAAYWGGTSMATPIYAGVNAOOLREHEFVNKRGITPRPSLLKALLAGAADKGLG	300
QY	507	PNGNGMGWRVLLDKSLNAAVYNNESSXLSTOSKATYXTATAGPKLISLWSDAPASTTA	566
Db	301	PNGNGMGWRVLLDKSLNAAVYNNESSXLSTOSKATYXTATAGPKLISLWSDAPASTTA	360
QY	567	SVTLVNDLDLVITAPNGTXYYGNDEFXAPXXXNMDGRNNVEVFINAXQSGTITIEVOAYN	626
Db	361	SVTLVNDLDLVITAPNGTXYYGNDEFTSEYNMDGRNNVEVFINAPOSQTYTIEVOAYN	420
QY	627	VPVGGQXESLAIYN 640	
Db	421	VPVGGQTESLAIYN 434	

### RESULT 3

ID	Q9A0R0.	PRELIMINARY;	PRT;	434 AA.
AC	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DE	Protease (Fragment).			
GN	PROE.			
OS	Bacillus sp. NV1.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;			
OC	Bacillaceae; Bacillus.			
OX	NCBI_TaxID=133781;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
SC	STRAIN-NV1;			
RX	MEDLINE=20568675; PubMed=11118284;			
RA	Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,			
RA	Horikoshi K.;			
RT	"Novel oxidatively stable subtilisin-like serine proteases from			
RT	alkaliphilic Bacillus spp.: enzymatic properties, sequences, and			
RT	evolutionary relationships.";			
RL	Biochem. Biophys. Res. Commun. 279:313-319(2000).			
DR	EMBL; AB046406; BAB2169.1; -.			
DR	HSSP; P00782; 1SUP.			
DR	InterPro; IPR000209; Peptidase_S8.			
DR	Pfam; PF00082; Peptidase_S8; 2.			
DR	PRINTS; PR00723; SUBTILISIN.			
DR	PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.			
DR	PROSITE; PS00138; SUBTILASE_SER; 1.			
FT	NON_TER 1			
FT	NON_TER 434 1			
SEQ	SEQUENCE 434 AA; 45294 MW; 7125D2 CRC64;			

Query Match	67.1%	Score 2082	DB 2	Length 434
Best Local Similarity	91.5%	Pred. No. 5.2e-118		
Matches 397; Conservative	14;	Mismatches 23;	Indels 0;	Gaps 0;

QY	207	NDVAVGIKAAVAAOSVYGLYQOGQIVAVADPGLDTGRDSSMBHAFGKLTALVALERTN	266
Db	1	NDVAVGIKAAVAAOSVYGLYQOGQIVAVADPGLDTGRDSSMBHAFGKLTALVALERTN	60
QY	267	NANDNGHGTHVAVSLVNGXNTKNGMAQOALVPOSLIMDSXGGLGGLPSMLQOTLEFSQAS	326
Db	61	NANDNGHGTHVAVSLVNGXNTKNGMAQOALVPOSLIMDSXGGLGGLPSMLQOTLEFSQAS	1200
QY	327	AGARIHTNSWGAANVAYATTDSRVAVDYVYKRNDMFTILFPAAGNEXPNGGTTSAEGTAKNAI	386
Db	121	AGARIHTNSWGAANVAYATTDSRVAVDYVYKRNDMFTILFPAAGNEXPNGGTTSAEGTAKNAI	180

```
QY 387 TVGATEMLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVAPGTXTILSARSSLPADSSF 446
|||||
DB 181 TVGATEMLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVAPGTXTILSARSSLPADSSF 240
|||||
QY 447 WANHDSKYAAYMGCTSMATPIVAGNVAQOLREHFVNKRGITPRPSILKAALIAAGAXGLGY 506
|||||
DB 241 WANHDSKYAAYMGCTSMATPIVAGNVAQOLREHFVNKRGITPRPSILKAALIAAGAXGLGY 300
|||||
QY 507 PNGNGGMRGTLDKSLNVAAYVNESSXLTSTOKATYXFTATAGKPLKISLWSDAPASTA 566
|||||
DB 301 PNGNGGMRGTLDKSLNVAAYVNESSXLTSTOKATYXFTATAGKPLKISLWSDAPASTA 360
|||||
QY 567 SYTLVNDLDLYTAPNGTXYVGNDEFXPPXXMWDGRNNVENVFINXQSGTYTIEVOAYN 626
|||||
DB 361 SYTLVNDLDLYTAPNGTXYVGNDEFXPPXXMWDGRNNVENVFINXQSGTYTIEVOAYN 420
|||||
QY 627 VPGPQXPSLAIVN 640
|||||
DB 421 VPGPQXPSLAIVN 434
|||||
```

## RESULT 4

```
Q9AOR1 PRELIMINARY: PRT: 433 AA.
ID Q9AOR1
AC Q9AOR4
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Protease (Fragment).
GN PROD.
OS Bacillus sp. SD521.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=13780;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SD521.
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL: AB046405; BAB21268.1; -.
DR HSSP: Q45670; IDBI.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 2.
DR PRINTS: PR00123; SUBTILISIN.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
FT NON-TER
FT NON-TER
SQ SEQUENCE 433 AA; 45576 MW; 98A2DF18FE660DDC CRC64;
```

```
Query Match 62.9%; Score 1952.5; DB 2; Length 433;
Best Local Similarity 86.6%; Pred. No. 3.4e-110;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;
QY 207 NDVARGIVKADVAAOSSYGLYGOGQIVAAVADTGLDTRNDSSMHEAFRGKITALVALGRTN 266
|||||
DB 1 NDVARGIVKADVAAONNNGLYGOGQVAAVADTGLDTRNDSSMHEAFRGKITALVALAGRNT 60
|||||
QY 267 NANPDMNGHGTIVAGSYLVGN-ALNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFQSAKS 326
|||||
DB 61 NANPDMNGHGTIVAGSYLVGN-ALNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFQSAKN 119
|||||
QY 327 AGARHTHTNSMGCAAYNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNGCTISAPGTAKNAI 386
|||||
DB 120 AGARHTHTNSMGCAAYNGAYTTANSRQVDEYVRKNDMTVILFAAGNEXPNGCTISAPGTAKNAI 179
|||||
QY 387 TVGATEMLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVAPGTXTILSARSSLPADSSF 446
|||||
```

```
DB 180 TVGATEMLRPSFGSLADNPNNHIAQFSSRGATRDGRIRKPDVATAGTFTILSARSSLPADSSF 239
|||||
QY 447 WANHDSKYAAYMGCTSMATPIVAGNVAQOLREHFVNKRGITPRPSILKAALIAAGAXGLGY 506
|||||
DB 240 WANHDSKYAAYMGCTSMATPIVAGNVAQOLREHFVNKRGITPRPSILKAALIAAGAXGLGY 299
|||||
QY 507 PNGNGGMRGTLDKSLNVAAYVNESSXLTSTOKATYXFTATAGKPLKISLWSDAPASTA 566
|||||
DB 300 PNGNGGMRGTLDKSLNVAAYVNESSXLTSTOKATYXFTATAGKPLKISLWSDAPASTA 359
|||||
QY 567 SYTLVNDLDLYTAPNGTXYVGNDEFXPPXXMWDGRNNVENVFINXQSGTYTIEVOAYN 626
|||||
DB 360 SYTLVNDLDLYTAPNGTXYVGNDEFXPPXXMWDGRNNVENVFINXQSGTYTIEVOAYN 419
|||||
QY 627 VPGPQXPSLAIVN 640
|||||
DB 420 VPGPQXPSLAIVN 433
|||||
```

## RESULT 5

```
Q9AOR4 PRELIMINARY: PRT: 433 AA.
ID Q9AOR4
AC Q9AOR4
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Protease (Fragment).
GN PROD.
OS Bacillus sp. D6.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=127889;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D6.
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL: AB046402; BAB21265.1; -.
DR HSSP: Q45670; IDBI.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 2.
DR PRINTS: PR00123; SUBTILISIN.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
FT NON-TER
FT NON-TER
SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;
```

```
Query Match 62.8%; Score 1948.5; DB 2; Length 433;
Best Local Similarity 86.4%; Pred. No. 5.9e-110;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;
QY 207 NDVARGIVKADVAAOSSYGLYGOGQIVAAVADTGLDTRNDSSMHEAFRGKITALVALGRTN 266
|||||
DB 1 NDVARGIVKADVAAONNNGLYGOGQVAAVADTGLDTRNDSSMHEAFRGKITALVALAGRNT 60
|||||
QY 267 NANPDMNGHGTIVAGSYLVGN-ALNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFQSAKS 326
|||||
DB 61 NANPDMNGHGTIVAGSYLVGN-ALNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFQSAKN 119
|||||
QY 327 AGARHTHTNSMGCAAYNGAYTTDSRNVDDYVRKNDMTILFAAGNEXPNGCTISAPGTAKNAI 386
|||||
DB 120 AGARHTHTNSMGCAAYNGAYTTANSRQVDEYVRKNDMTVILFAAGNEXPNGCTISAPGTAKNAI 179
|||||
QY 387 TVGATEMLRPSFGSYADNINHVAQFSSRGPTKGRIRKPDVAPGTXTILSARSSLPADSSF 446
|||||
DB 180 TVGATEMLRPSFGSIADNPNNHIAQFSSRGATRDGRIRKPDVATAGTFTILSARSSLPADSSF 239
|||||
```

QY	447	WANHSKAYVGGGSMAPPIYAGVAVAOOLREHFVKNRGITTPSPSLTKKALLAGAADXGLGY	5066
Db	240	WANHSKAYVGGGSMAPPIYAGVAVAOOLREHFVKNRGITTPSPSLTKKALLAGAADVGLGY	2999
QY	507	PNGNCGMGRVTLDSLSNAVYNNESSXLSTSOQATYTXFTATAGKPLKLSLWSDAPASTTA	5666
Db	300	PSGDGWMGRVTLDSLSNAVYNNEAFALTITGGKATYSPQTQAGKPLKLSLWTDAPGSTTA	3599
QY	567	SVTLVNDLDLVTAPNGYXYYGNDPFXAXXXXNWDGRNNVENVFINXQSGTYTTEVOAYN	6288
Db	360	SVTLVNDLDLVTAPNGKYVGNDFSPYDNNMDGRNNVENVFINAQSGTYTTEVOAYN	4199
QY	627	VVPVQGFESLAIYN 640	
Db	420	VPSQGFRESLAIYH 433	
RESULT 6			
Q9A0R2			
ID	Q9A0R2	PRELIMINARY;	PRF; 433 AA.
AC	Q9A0R2	01-JUN-2001 (TREMBlrel. 17, Created)	
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)		
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)		
DE	Protease (Fragment).		
GN	PROC.		
OS	Bacillus sp. Y.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Bacillaceae; Bacillus.		
OX	NCHI_TaxID=133779;		
RM	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-Y.		
RX	MEDLINE=20566675; PubMed=11118284;		
RA	Saeiki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,		
RA	Horikoshi K.		
RT	"Novel oxidatively stable subtilisin-like serine proteases from		
RT	alkaliphilic Bacillus spp.: enzymatic properties, sequences, and		
RT	evolutionary relationships."		
RL	Biochem. Biophys. Res. Commun. 279:313-319(2000).		
DR	EMBL; AB046404; BAB21267.1; -.		
DR	HSSP; Q45670; IDBI.		
DR	InterPro; IPR000209; Peptidase_S8.		
DR	Pfam; PF00082; Peptidase_S8; 2.		
DR	PRINTS; PR00723; SUBTILISIN.		
DR	PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.		
DR	PROSITE; PS00136; SUBTILASE_SER; 1.		
FT	NON_TER	1	
FT	NON_TER	433	
FT	SEQUENCE	433 AA; 45587 MW; B01291A803C775AE CRC64;	
Query Match			
Best local similarity 62.6%; Score 1941.5; DB 2; Length 433;			
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps			
QY	207	NDVARGIKADVAVOOSYVGLGQGIIVAVADTGLDTGRNDSMHEAFGKITAYALGRTN	2666
Db	1	NDVARGIKADVAVOONNGVLGYQGQGVAVADTGLDTGRNDSMHEAFGKITAYALGRTN	60
QY	267	NANDTNGGHTHVAGSVLGGXKTKNGMAPOANLVFOQIMDSXKGLGLPSNLQTLFSGAXS	3288
Db	61	NASDNGGHTHVAGSVLGN-ALNKGMAPOANLVFOQIMDSXKGLGLPSNLTLFSGAWN	1199
QY	327	AGARHTHSWGAANVAGATTDSRNVDDVVRKNDMTTLFAAGNEKPNNGTISAPGTANAI	3866
Db	120	AGARHTHSWGAANVAGATTANSKRDVDEVRNNDMTTVLFAAGNEKPNNGTISAPGTANAI	1799
QY	387	TVGATENLRPSFGSYADININVAOFFSSKGPDKDIRIPDVAWPEPTXILISARSLAPDSF	4466
Db	180	TVGATENLRPSFGSIAIDPNHIAQFSSSGAIRDORIRKPDVTAPPTFILSARSLAPDSF	2399
QY	447	WANHDSKAYVGGGSMAPPIYAGVAVAOOLREHFVKNRGITTPSPSLTKKALLAGAADXGLGY	5066

Dd	240	WANTSKAYAMGCTSMARPIYAGNAAQLREHRIKNGITPKRSLIKALINAGATDVGLGY	299
Oy	507	PNGNDGWGRVTLDKSLNAAYVNESXLSSTSQKATYXFPTATACKPLKISLWSDAPOSTTA	566
Dd	300	PNGDQGMGRVRVLTKNSLINAAYVEAFALATGCKATYSFOQAOKPULKISLWMDAQSFTA	359
Oy	567	SVTLLVNDLDLVTTANGXYVCNDPXXPXKXMMWDNRNNVENFIMXPSCGTTEIVOAQN	626
Dd	360	SYTLLVNDLDLVTTANGKYVGNDFTYPYDNMWRNNVENFVINAPOSITYIIEVOAYN	419
Oy	627	VPGVGPQXSFLATVFN 640	
Dd	420	VPSGPQRRESLATVH 433	
 RESULT 7 O8T9W1 PRELIMITARY: PRT: 1825 AA.			
AC	O8T9W1:		
DT	01-JUN-2002 (TREMBLrel. 21, Created)		
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)		
DE	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DR	Serine protease/ABC transporter Tagd.		
CN	TAGD.		
OS	Dictyostellium discoideum (Slime mold).		
OC	Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium,		
OX	NCHI_TaxID=44689;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=AX;		
RA	Anjaard C., Loomis W.F.;		
RT	"Evolution of the ABC transporters of Dictyostellium."		
RL	Submitted (JAN-2002) to the EMBL/Genebank/DBJ databases.		
DR	EMBL: AF46309; AAL/4253.1; -		
KW	Protease.		
SO	SEQUENCE 1825 AA; 202641 MW; E28160BC7613A3B CRC64;		
 Query Match 15.9%; Score 493; DB 5; Length 1825; Best Local Similarity 24.2%; Pred. No. 2.3e-21; Matches 192; Conservative 105; Mismatches 229; Indels 268; Gaps 31;			
Oy	88	HIXQNGPLEET---KQLEXTG-----AKIIDYTPDAYIV---EYEDGVXSX 131	
Dd	127	YIVOGRKHINDETRPOFKOFLINTDIVLDEOPYQSHINYVIPHSDFVLMNDEQSNLLSS 186	
Oy	132	XXXIHVESVP-----YLPPXYXDIPOLFKGASXL-----VK 164	
Dd	167	KEWVSWIGFEPEFSNKILHLNYENKSIGLPYITKLSTNSLIORMENTLSILTYSNVK 246	
Oy	165	AxAADTKQXKNKEVOLRGIEIXIAQXXS-----NDVXYITAKEPYRVMDV 209	
Dd	247	LTLINQKTL-KSIYVCNDESSOSSCSLYVSSEKLYOWMTSEQSESNTYERSEKFOTANRL 305	
Oy	210	ARGIV-----KADVAOSSIGLYGGCQIVAVADTGIDTGR--NDS-----SM 248	
Dd	306	SPKAIFGTKRTLNVNDRRIDIP---LRKGQILSIADTGLDGSHCFSDSNNPPIPNVSV 360	
Oy	249	HEAFGKTITALYALGRTNNANDNGTHGHVAGSVGNGXTN-----KGMAPQANLVFO 301	
Dd	361	NLNHRKVYTIISL-HDNEDVIDOHGHVCGSAAGAEPDSSLSSFGSLATDKAIFAE 418	
Oy	302	SI-MDSXGGLGPSNLQTLFESQASAGARIHTNSWGA---AVNGAYTTDSRNVDYVR 356	
Dd	419	DLASPSPNNNEPVPPEDYSOLYQPLYNAGARVHGDSMSLSIGYLGYSDDAGSIDDFLY 478	
Oy	357	KN-DMTFIIPAAGNEKPNGCTISAPCTAKNATVGTENLRPSF----- 398	
Dd	479	THPDFIIIRAGGNNOYSSLLS-QATAKNVITVGAEGTTHESYTDALEYSNFETVAKST 537	
Oy	399	GSYAD-----NTNHQAQ 410	
Dd	538	INSLCQSFDDDKCYTYTAQCCEYSTVKGLSGCCSTSYIKNSNJAASFSSOPELYNENNTICS 597	



```

OY 174 NKEVOLRGIEIXIAOXXXXNDVXYITAKPEYKVMNDVARGIVKADY----- 218
DB 86 NKPVSDADLFTTAK-----NICKFNKHKRYKTIIPSTIAANULTSOIWNLSKLEIVKQIEND 140
OY 219 -----AOSSYGLYGOQOIVAADVADTGLDTRNDSSMEHAEFRGKITALLY 260
DB 141 EPPVATLDATKMFCTKARSDFVGKNTITAIIDTGDIGNHVDLS-----GKRT----- 191
OY 261 AAGRTNNAN-----DTNGHGHVAGSVLGNXTN---KGMPQANLVFQSIIMDS--XG 309
DB 192 -IGMKDFINNTKTPPYDDNGHGHVVASIAAGTAGNSEFYGVAPDALVGIKVLDAANGSS 250
OY 310 LGLPSNIQTLFSAAXSAGARIFHTNSMGAANVAGATTDSDRNDVYRKNDMTILFAGNE 369
DB 251 MGTVAGIDMAVONNDVGIKININLSGTSSDGDSTSLAVNRADVGIYVVAAGNS 310
OY 370 XNPGGTISAPGTAKNAITVGATENLRPSFGYADININHYAOFSSRGPTKGRKIDPVAP 429
DB 311 GPAKYTIGSPGAAEKAITVAAADY---GELGFNL---ASFSSRGPTADGRKIDPIAP 363
OY 430 GTXILSARSLAPDSSSFANHDSKAYMGCTSMATPIVAGNVAOLREHFVKRGTIPKRS 489
DB 364 GNTITPAK-----ANSVNGYVYSGTSMATPEVAGFVALMLN---ANPNLTPNDA 410
OY 490 LTKAALIGAADXLGYPNGNOGMRVTLDKSLNVA-----YNNESSXLST 535
DB 411 --KNIMSTAKSMGPPSKNVQDYGAGRLDGEALRVAAGNRGNNDIVPNHYI--SGYLP 466
OY 536 SOKA-TYXHTAT-AGKPKLISLV---WSDAPASTASVTLVNDLVTAPNGTXYXVND 590
DB 467 SKYSDTWTFTNANTSTYPIAITLIPDMANYNP-----DFDIYLPDSTGLKSS- 515
OY 591 FXXPXXNMDDGRNNVENVFINKPOSGTYTIEVOAY 625
DB 516 -----TGTORETITILPSGTGYVVKYSY 541

RESULT 10
OY 08UOC9 PRELIMINARY: PRT: 654 AA.
AC 08UOC9:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Alkaline serine protease.
GN PF1670.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RT Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.:
  "The complete sequence of the Pyrococcus furiosus genome.";
  Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AE010265; AAL81794.1; -.
KW Protease; Complete proteome.
SQ SEQUENCE 654 AA: 70230 MW: 1CB145A5F505DB34 CRC64;

```

Query Match: 13.2%; Score 408; DB 17; Length 654;  
 Best Local Similarity 25.9%; Pred. No. 8,6e-17;  
 Matches 162; Conservative 80; Mismatches 210; Indels 174; Gaps 26;

```

OY 63 VESENKXKLGKAKKLETPANKKIH--IXQNGPLLETKOXLETKKIIDYIDPVAY 120
DB 35 VERNKGLPLPGFRKIQKINPNEISTYVFNHREKELAVVLELMGAKV-----RY 87
OY 121 IVEYEGDVXXXXXIEHVESVEPLYPYXIDPOLFTKGSXLYKXALDTQXNKEVOLR 180
DB 88 V-----YHITPAI---AADLKVRDLIVISGLTGKAKLS 118

```

```

OY 181 GIEIXIAOXXXXNDVXYITAKPEYKVMNDVARGIVKADYAOSSSYGLYGOQOIVAADVADTGLD 240
DB 119 GVRFIQEDYK-----VYVSALBEGDESAQVMAFYWMNIGYD--GSGITIGIIDTGID 170
OY 241 TGRNDSSMEHAEFRGKITALLYALGRTNNAN-----DTNGHGHVAGSVLGNXTN-----K 290
DB 171 -----ASHPDLOGRV-----IGWDFVNGRSYPYDDHGHGHVVASIAAGTAGAANGRYK 219
OY 291 GMAPQANLVFQSI--DSXGLGLPSNIQTLFSAAXSAGARIFHTNSMGA----- 338
DB 220 GMAPQAKLAGIKVVLGADSGSISTITIKGYEMAVDNKDKYIKVIMLSGSSQSSDGTAL 279
OY 339 --AVNGAYTTDSRNDVYRKNDMTILFPAENGXPNGGTISAPGTAKNAITVGATENLRP 396
DB 280 SQAVNAAMD-----GLVVVAAAGSGPNKTYIGSPAASKVITVGA----- 321
OY 397 SFGSYADININHYAOFSSRGPTKGRKIDPVAPGTXILSARSLAPDSSSFANHDSKAY 456
DB 322 -----VDKDYVITSPSSRGPADGRKLEVPAPGMMIIAARAS---GTSMGOPINDYTTA 373
OY 457 MGTSMATPIVAGNVAOLREHFVKRGTIPK--PSLKAALITGA-----ADXLGLY 506
DB 374 APTSMATPIHAGIALLLQ-----AHPSTPDKVKTALLETADYKPEIAD--IAY 424
OY 507 PNGNOGMRVTLDKSLNVAAYNESSXLSTSOXA---TYXETAAGRPKLISLVMSDAP 561
DB 425 -----GAGRVNAYKAIN--YDNVAKLVEFGYVANKGSGQHGVISGASFVETALYMDAN 477
OY 562 ASTTASVTLVNDLVTAPNG-----TXVGDGDFXPPXXNMDDGRNNVENVFINKPO 614
DB 478 -----SDLDLYLVDPNNGOVDSYTAIY-----FEKVGYINPT 511
OY 615 SGTYTEVOAYNVVPQXFSIAVIN 640
DB 512 DGTWITIKVSYSS--GSANYQYDVVS 534

```

```

RESULT 11
OY 09FBZ4 PRELIMINARY: PRT: 1239 AA.
AC 09FBZ4:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative secreted peptidase.
GN SC07188 OR SC8A11.16C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Saunders D.C., Harris D.;
  Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
  Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
  Kinashi H., Hopwood D.A.;
  "A set of ordered cosmids and a detailed genetic and physical map for
  the 8 Mb streptomycetes coelicolor A3(2) chromosome.";
  Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,

```



RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT \*Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RT Nature 417:141-147(2002).  
 DR EMBL: AL391041; CAC01588.1; -  
 DR HSSP: 099405; IMPT.  
 DR InterPro: IPR000317; PA.  
 DR InterPro: IPR000209; Peptidase\_S8.  
 DR Pfam: PF00225; PA; 1.  
 DR Pfam: PF00082; Peptidase\_S8; 1.  
 DR PRINTS: PR00723; SUBTILISIN.  
 DR PROSITE: PS00136; SUBTILASE\_ASP; UNKNOWN\_1.  
 DR PROSITE: PS00137; SUBTILASE\_HIS; UNKNOWN\_1.  
 DR PROSITE: PS00138; SUBTILASE\_SER; 1.  
 SO SEQUENCE 1239 AA; 128505 MW; 8F5E9AC68EB1260A CRC64;

Query Match 12.6%; Score 391; DB 16; Length 1239;  
 Best Local Similarity 30.9%; Pred. No. 2e-15;  
 Matches 143; Conservative 51; Mismatches 187; Indels 82; Gaps 16;

QY 214 VKADVAOSSY-----CLYGQGIYAVADTGLDGRDSSMHEAFRGKITALXALG 263  
 DB 219 VEADLADSTAGIAPRAMAGNTGGVEAVLDITGDAG-----HPDLADRIAAQSY 272  
 QY 264 KTNNAADTNGGTHVAGSYLGKXNTN---KMAPOANLVQSIIMDSXGLGLPSNLOT 319  
 DB 273 PRENTDDGDGHTVASTINGTGAAGKAGKAGARLSTGKVLDN--SRGQISMTLAA 331  
 QY 320 LFSQAXSAGARIHTNSGAAVNCAYTDD--SRVVDYVRKNDMTILFAAGEXNGGTISA 378  
 DB 332 MEMAAVERHAATVMSLGSSEQSDGSDPSRAVDRLSAQTALFVVAAG--GCEAGSISA 390  
 QY 379 PGTAANATVVGATENLRPSFGSYADININVAOFSSRCPTKDKRIKIPVMAPGTYLSARS 438  
 DB 391 PEVATSAITVGA-----VDATDILAPFSSQGFVDAKLPETAPGVGILAA-- 437  
 QY 439 SLAPSSFWANHDSKYAVMGTSMAPIYAGNVAQLREHFVKNNGITPKPSLLKALV-- 495  
 DB 438 -----NSSFAAGNGAYOSLSTGSMATPHVAGAAALL-----AARPDLSGSLKDY 484  
 QY 496 IAGADXLGYPNGNOGMRVTLDKSLN-----VAYVNESSXLSISQKATYFTRATAG 548  
 DB 485 LASSSHRPRIDAFQAGSGRVDAAVRAGVYASATAPGSSFGPVRLVYTTNTTGA 544  
 QY 549 KFLKISLWMSA-----PASTASVTLVNDLDTLITAPNGTXYVGNDFXXP 594  
 DB 545 VTLELSVATHTAPGCVFRLSASRYVPAHGTADVTLLTIDGS--GSAGGRAVSGQILATD 601  
 QY 595 XXXXNDGNVENVFINKPQSGTYTLEVO--AYNVVGFQFXSL 636  
 DB 602 A-----DARNVAHTAVSAGPVHRKLTIVHFKDADGNPV--PGVEFL 639

RESULT 12  
 Q9FC06 PRELIMINARY; PRT: 1253 AA.

AC 09FC06;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative secreted peptidase.  
 GN SC0176 OR SC8A11.04C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Filicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OC NCBI\_TaxID=1502;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Saunders D.C., Harris D.;  
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA MEDLINE-97000351; Pubmed-8843436;  
 RA Kienast H., Hopwood D.A.;  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT \*Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RT Nature 417:141-147(2002).  
 DR EMBL: AL391041; CAC01576.1; -  
 DR HSSP: 099405; IMPT.  
 DR InterPro: IPR003137; PA.  
 DR InterPro: IPR000209; Peptidase\_S8.  
 DR Pfam: PF00225; PA; 1.  
 DR Pfam: PF00082; Peptidase\_S8; 1.  
 DR PRINTS: PR00723; SUBTILISIN.  
 DR PROSITE: PS00136; SUBTILASE\_ASP; UNKNOWN\_1.  
 DR PROSITE: PS00137; SUBTILASE\_HIS; UNKNOWN\_1.  
 DR PROSITE: PS00138; SUBTILASE\_SER; 1.  
 SO SEQUENCE 1253 AA; 130971 MW; AA69BA17FEEDB89 CRC64;

Query Match 11.9%; Score 368.5; DB 16; Length 1253;  
 Best Local Similarity 28.0%; Pred. No. 4.7e-14;  
 Matches 142; Conservative 55; Mismatches 169; Indels 141; Gaps 20;

QY 206 MNDVAGTIVKADVAOSSYGLYGQGIYAVADTGLDGRDSSMHEAFRGKITALVAGRT 265  
 DB 219 ISDTTAQIGAPDV--MSGNGTGGGVAVAVLDITGDAG-----HPDFAGRIATAVSFVPD 270  
 QY 266 NNANDTNGGTHVAGSYLGKXNTN---KMAPOANLVQSIIMDSXG----- 308  
 DB 271 QDVTRNGGHTVASTVAGTGAAGSGVEKGVAPGASLHIGKVLNDSGSDGQSVLAGMEM 330  
 QY 309 -----GLGGLPSNLQTLFQAXSAGARIHTNSMGAANVNCAYTDDSNVNDYVR 356  
 DB 331 AVROOHAKTIVMSLGDSPDTGDTPLSEAV-----NMLSATGCA----- 368  
 QY 357 KNDMTILFAAGNEXPNCGTISAGTAKNATVVGATENLRPSFGSYADININVAOFSSRCP 416  
 DB 369 -----LFVVAAGNSCFEATYVGTTPAADAAVLGAVNG--PEKG-----VQDLADSSSKCP 417  
 QY 417 -TKDGRIRPDVMAPGTYILSARSLAPDSSFWANHDSKYAVMGTSMAPIYAGNVAQLR 475  
 DB 418 RYGDNAVAPDLTAPGCVGLAARSYAPBEG-----EGAYOSLSTGSMATPHVAGAAALLA 471  
 QY 476 EHFVKNRGTTPKPSLKAALAGADXLGYPNGNOGMRVTLDKSLNVAAYVNESSXLSLT 535  
 DB 472 AEHPDMTG-----ORLKEALVGTGTAGTORFSP--FDAGSGRV-----DVAAAVRSTLLAS 519

```
QY      536 SQ--KATY-----XFLPATAGRKLPLKISLVMSDA-----PAST 564
           :| | :| | :| | :| | :| | :| | :| | :| |
Db      520 GDAFAQAATPYPTPGQTVRBDVITYNSGPAPALDIALSPALEPGCLFTLSEAQTYVAHG 579
           :| | :| | :| | :| | :| | :| | :| |
QY      565 TASYTLVNDLD-----LVITAPNGXYVGNDFFXAPXXXNMCGRRNVVEVF--- 610
           |||| :| | :| | :| | :| | :| | :| | :| |
Db      580 TASVGVITTHDLAEDNGAYATRLVASGADGAVLA----RPFGVKGKGRATLALTAKDH 635
           :| | :| | :| | :| | :| | :| | :| |
QY      611 -NXPSGGTYTIEVQAYNVPVPOXESL 636
           :| | | | :| | :| | :| | :| |
Db      636 HDKPLSGTVILTKDYERN-TAPKVYSV 660
```

## RESULT 13

ID	PRELIMINARY:	PR:	1102 AA.
AC	P95684;		
DT	01-MAY-1997 (TREMBLrel. 03, Created)		
DT	01-MAY-1997 (TREMBLrel. 03, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Subtilisin-like protease.		
OS	Streptomyces alboborisolus.		
OC	Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae;		
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces		
OX	NCBI_TaxID=1887;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRATN-S-3253;		
RX	MEDLINE=97144528; PubMed=8990295;		
RA	Suzuki M., Taguchi S., Yamada S., Kojima S., Miura K., Momose H.;		
RT	A novel member of the subtilisin-like protease family from		
RT	Streptomyces alboborisolus. "		
RL	J. Bacteriol. 179:430-438(1997).		
DR	EMBL: D83672; BAA12040.1; "		
DR	HSSP: P00782; 2S8T.		
DR	InterPro: IPR002860; GH_BNR.		
DR	InterPro: IPR000209; Peptidase_S8.		
DR	Pfam: PF02012; BNR: 2.		
DR	Pfam: PF00082; Peptidase_S8; 1.		
DR	PRINTS: PRO0723; SUBTILISIN.		
DR	PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.		
DR	PROSITE: PS00137; SUBTILASE_HIS; 1.		
DR	PROSITE: PS00138; SUBTILASE_SBR; 1.		
DR	Protease.		
QO	SEQUENCE	1102 AA; 114128 MW; F9E4AD250FE55E CRC64;	

```

Db 433 -----WKRAELKGLTASTKNC-----KYHPEEGSSRNOVDKAIQTVIAEVSLSFGQOM 505
QY 535 -----TSQATYXFTYTAACKPLKISLWMD-----APAS-----TTSATVELNDLDLVITA 580
Db 506 PHADDKPVYTKLLTIRNIGEDTYTLTKLSTAGCPKGAAPFFLGASL-----IV 537
QY 581 P-NGTYGVNDPXXPXXNMWGRNNVENVFINXQS-----GYTIEVQAYNV 627
Db 558 PANGTASVDYATADTRLGAGVGTSAVVAVATGAQOSVRTAAVEREVESYNV 609

```

## RESULT 14

ID	PRELIMINARY:	PRT:	444 AA.
09KBJ7:			
DT 01-OCT-2000 (TREMblrel. 15, Created)			
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)			
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)			
DE Intracellular alkaline serine protease.			
GN APRX OR BH1930.			
OS <i>Bacillus halodurans</i> .			
OC Bacteria, Firmicutes; Bacillus/Clostridium group; Bacillales;			
OC Bacillaceae; Bacillus.			
OX NCBI_TaxID=86665;			
[1]			
SEQUENCE FROM N.A.			
RP STRAIN=C-125 / JCM 9153;			
RX MEDLINE=20512582; PubMed=11058132;			
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,			
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kohara S.,			
RA Horikoshi K.;			
RT "Complete genome sequence of the alkaliphilic bacterium <i>Bacillus</i>			
RT <i>halodurans</i> and genomic sequence comparison with <i>Bacillus subtilis</i> ."			
RL Nucleic Acids Res. 28:4317-4331(2000).			
DR EMBL: G0901513; BAB05649.1; -			
DR HSSP: Q99405; IMP.			
DR InterPro: IPR001128; Cytochrome_P450.			
DR InterPro: IPR000209; Peptidase_S8.			
DR Pfam: PF00082; Peptidase_S8; 1.			
DR PRINTS: PR00723; SUBTILISIN.			
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.			
DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.			
DR PROSITE: PS00137; SUBTILASE_HTS; 1.			
DR PROSITE: PS00138; SUBTILASE_SER; 1.			
DR Protease; Complete proteome.			
Q0 SEQUENCE 444 AA; 3B05F8C53A21F2F CRC64;			

DB 371 RVGSHYTMNSGISMATPVCAGVALMLQH---EPNLTPDE--VKTRLMEGTDRMADRDPN 425  
QY 502 -XGLGY-----PNCNQ 511  
DB 426 VYGAGYISAEGAIPNSE 443  
RESULT 15  
O9RL54  
ID O9RL54 PRELIMINARY: PRT: 1245 AA.  
AC O9RL54  
DT 01-MAY-2000 (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, last sequence update)  
DT 01-JUN-2002 (TREMREL. 21, last annotation update)  
DE Probable secreted peptidase.  
CN SC00432 OR SCF51A.10  
OS Streptomyces coelicolor  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_taxid:1902;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Murphy L., Harris D.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA MEDLINE:97000351; PubMed:8843436;  
RA Redenbach M., Kleser H.M., Denapante D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmid and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
[4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2) / M145;  
RA Bentley S.D., Chaler K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,  
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete Streptomyces  
coelicolor A3(2).";  
RL Nature 417:441-447(2002).  
DB EMBL; ALI21596; CAB56662.1; -  
DR HSP; P00782; ISOE.  
DR InterPro: IPR003137; PA.  
DR InterPro: IPR000209; Peptidase\_S8.  
DR Pfam: PF02225; PA; 1.  
DR Pfam: PF00082; Peptidase\_S8; 1.  
DR PRINTS: PR00723; SUBTILISIN.  
DR PROSITE: PS00137; SUBTILASE\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00138; SUBTILASE\_SER; 1.  
SQ SEQUENCE 1245 AA; 130895 MW; 74EE92DB9CAIDEGO CRC64;  
Query Match 10.9%; Score 337; DB 16; Length 1245;  
Best Local Similarity 31.5%; Pred. No. 3,7e-12;  
Matches 123; Conservative 40; Mismatches 156; Indels 70; Gaps 14;  
QY 222 SYGLXGGQIVANADTGLDTRGNSMHEARFKITATLALGRTNNANDTNGGTHVA-- 279  
DB 227 SAGTRGDGVAAVADLTGAD-----QSHPDLAGVAAAKDFSGSGTNDVFGHGTHTVASI 280

QY 280 -GSVLNGXTNKGAPQANLVFQSIINDSXGGLGLP.SNLIQTLFQOAXSAGARITHNSWG 337  
DB 281 VGGSGAASGSRQGVAAFAARLLVGKVLGDG--FGSESOYINGMEMADQGADAVVNMSLG 338  
QY 338 A--AVNGAYTTD--SRNVDDYVRKNDMTILFAAGNEXPNG-GTISAPGTAKNAITVGATE 392  
DB 339 SSGATDG---TDPMSQALINDLSRRRTGLFVYAAGNNEGEGPRTVSGPAAADALTVGA-- 393  
QY 393 NLRPFGSYADNINVAQFSSRGP-TKDGRKPDVAPGTXILISARSSLAPDSSFMANH 451  
DB 394 -----VDRDSDLAPFSSRGRPLGDDAVKPDYATPGVGIAR---AAGSAMGDPVD 441  
QY 452 SKYAYMGSTMAPTIYAGNVAQLREHVKKRGITPKPKLLKALLAGAADGLCPNCQ 511  
DB 442 EHTTAASGSMATPHVAGAAALLQRRPDMTGAQLDALISTRAIVD-----GQKYTEQ 495  
QY 512 GWGRV-----TLKSLNVAIVNESSXLSSTOKATYKFTAT 546  
DB 496 GGGRIDVRAAGLGAATATGTLVWGPTTSRDTPEVTSRVRYTNSDSDVTLSLAVALATEG 555  
QY 547 AGRPKLSL-VMSDA--PASTTASVTLVND 573  
DB 556 GKAPAEASRLGSDSVRPAGSSAEVPLTVD 586

Search completed: April 1, 2003, 12:40:57  
Job time : 63.5473 secs

**THIS PAGE BLANK (USPTO)**